

M
线性代数引论
（第二版）
蓝以中 赵春来 编著
北京大学出版社

线性代数引论

责任编辑 王明舟
封面设计 张　虹



ISBN 7-301-03712-0/O · 417 定价: 16.50元

线性代数引论

（第二版）

蓝以中　赵春来　编著

北京大学出版社
北　京

图书在版编目(CIP)数据

线性代数引论/蓝以中，赵春来编著．—2版．—北京：北京大学出版社，1998.4

ISBN 7-301-03712-0

Ⅰ.线… Ⅱ.①蓝…②赵… Ⅲ.线性代数 Ⅳ.O151.2

书　　名：线性代数引论（第二版）
著作责任者：蓝以中 赵春来
责 任 编 辑：[illegible]
标 准 书 号：ISBN 7-301-03712-0/O·417
出　版　者：北京大学出版社
地　　址：北京大学校内 100871
电　　话：出版部 62752015 发行部 62559712 编辑部 62752032
排　印　者：北京大学印刷厂
发　行　者：北京大学出版社
经　销　者：新华书店
850×1168 32开本 12.875印张 320千字
1998年4月第一版 1998年4月第一次印刷
定　　价：16.50元

前　言

近年来，由于自然科学和工程技术的迅速发展，特别是由于电子计算机的普遍使用，使得线性代数得到日益广泛的应用，因而也就对综合大学和理工院校线性代数课的教学内容提出了越来越高的要求．通过线性代数课程的教学，除了让学生了解一些基本的线性代数计算问题之外，还应当进一步使他们对于线性代数的基础理论有较深的了解，以便融汇贯通地运用线性代数的工具去解决理论上和实践中遇到的各种问题．

在我国综合大学和理工院校的高等数学课程中，线性代数一般安排在微积分之后，因而学习线性代数时，学生已经有了一定程度的数学训练，但从教学实践中看，学生从微积分转到学习线性代数时，仍然感到比较困难．这其中一个主要的原因是线性代数中所研讨的 n 维向量、矩阵、线性空间、线性变换，与学生在前面课程中已经熟悉的实数及其代数运算差别很大，初学者往往感到难于理解．因此，一本较适用的线性代数教材，一方面应当具有足够的理论深度，以满足近代自然科学和工程技术的需要；另一方面，又应当尽可能由浅入深，使初学者感到入门并不难，从而提高深入掌握其基本理论和基本方法的信心．

本书就是根据以上的认识编写的．既包括了与综合大学数学系线性代数课程相同的内容，在阐述上也保持了数学专业教材所应有的逻辑上的严格性；同时，又较多地照顾到非数学专业的学生和非数学专业工作者的特点，对于线性代数基本概念的引入和基本命题的阐述尽可能详尽，由具体而抽象，力求使之通俗易懂，不使初学者望而生畏．在读者理解和掌握了线性代数的基本概念和方法之后，再逐步加深内容，使达到必要的理论高度．

本书共包含八章，其中带 * 号的章、节和段落是属于选讲的内容. 如果教学计划中线性代数课时可达 60 学时左右（不包括习题课），预计可以讲完前七章的全部内容. 如果仅有 50 学时左右，则可略去第七章和前六章中带 * 号的大部分节和段落. 而如果仅有 30～40 学时，则建议讲授第一、二章全部，第三章 §1 和 §2（略去行列式性质的证明），第四章 §1、§2 以及子空间的概念，第五章 §1，§2，§3，第六章 §1，§2.

本书删去了 λ-矩阵的理论. 因为对于非数学专业的学生，它用处不大而又难于掌握，勉强讲授往往事倍功半. 而对于数学专业的学生，这一部分内容又可以在将来的抽象代数课程中从更高的观点上得到简单明了的处理. 因此，本书对矩阵的若当标准形只作了叙述而略去有关定理的证明. 考虑到在实际工作中可能需要计算低阶矩阵的若当标准形，在第五章 §4 中添加了求二、三阶矩阵的若当标准形一段. 此段内容仅供参考，可不作为教学内容.

在本书中，定理按章编号，命题和例题按节编号. 在行文中，当需要跨章引用前面的命题或定理时，就指明是第几章的命题或定理. 如无说明，则所引用的均是该章的命题或定理. 一个命题或定理证明完毕时，以“▌”号表示之.

本书是以作者近年来在北京大学讲授线性代数时的讲稿为基础写成的. 在编写过程中，得到丁石孙同志的热情指导与帮助，他仔细地审阅了全书，提出了许多宝贵的修改意见，从而使本书的质量得以有很大的提高，作者在此向他表示衷心的感谢. 本书中吸收了北京大学数学系几何代数教研室代数小组所编《高等代数》一书中的许多习题，在此一并致谢.

由于作者水平的限制，书中不当和错误之处一定不少，诚恳地希望读者批评指正.

编　者

一九八一年元月于北京大学

第二版序言

本书自一九八一年出版以来，在北京大学理科大多数系被作为“线性代数”课程的教材使用，已经整整十五年了. 在这期间，又由台湾高等教育出版社于一九八九年在台湾出版了繁体字版本. 现在我们根据十五年来教学实践的经验，对此书作了一次全面的修订，目的是使它更加符合教学实践的要求.

在这篇序言中，首先要向读者说明的是：线性代数的研究对象是什么?根据现代数学的观点，代数学的研究对象是各种代数系统及其相互关系，即这些代数系统之间的同态映射. 线性代数是代数学中一个最初等的分支，它研究一类最简单的代数系统：线性空间及其同态映射，即线性映射，特别是线性变换. 对线性代数研究对象的这一观点，是统率本书总的指导思想. 从开篇第一章到全书终了，所有材料都围绕这一线索组织起来，按照由具体到抽象，由特殊到一般的原则，依次展开. 首先阐述具体的向量空间和矩阵的基础理论，在读者逐步适应了这些具体的研究对象之后，再过渡到抽象和一般性的线性空间、线性变换. 我们希望读者充分注意全书这一总的指导思想，以便学习时处于主动的地位，不致被书中众多的定义、命题所迷惑.

作为一本教材，特别是大学一、二年级基础课的教材，其理论的阐述，概念的引入，要符合人的认识规律. 代数学是一门较为抽象的学科. 多年教学经验证明，不适应或不理解线性代数中一些抽象概念，是学生学习中遇到的最大困难. 如何根据教育学的基本原则，把线性代数中各种抽象概念阐述的通俗易懂，使缺乏抽象数学训练的低年级学生能较好地接受这些概念，是我们编写这本书时着重考虑的一个问题. 在编写本书第一版时，我们已采取了一系列

措施来帮助学生克服学习中的这个难点.十五年来的实践证明,这些措施是行之有效的.这次修订时,我们又做了一些努力,希望它能成为一本更加适用的线性代数教科书.

最后,我们要对为本书初版和第二版付出大量辛勤劳动的北京大学出版社邱淑清和王明舟同志表示衷心的感谢.

编　　者
一九九六年十二月

目　录

第一章　线性方程组

引　　言

读者在中学的课程中已经熟悉了有关二元一次方程组和三元一次方程组的基本知识. 本章的任务，是要对中学里学过的这些知识从理论上加以总结和提高.

二元一次方程组的一般形式是

$$\begin{cases} a_1x + b_1y = c_1, \\ a_2x + b_2y = c_2. \end{cases} \tag{1}$$

在解析几何中我们已经知道：在平面直角坐标系内，每个二元一次方程代表一条直线. 由于这个原因，一次代数方程也称为线性方程. 今后，我们将未知量的一次方程统称为线性方程（不管未知量的数目有多少）.

读者已有的关于二元和三元线性方程组的知识大致有如下两个方面：

（1）方程组的求解方法. 在中学代数课程中已经指出：解二元和三元线性方程组的最基本的方法是消元法（代入消元法或加减消元法）. 例如，从二元线性方程组(1)中设法消去未知量 y，得到未知量 x 的一个一元一次方程，解出这个一元一次方程得到 x 的值，再代回原方程组求未知量 y 的值. 这个方法也适用于三元线性方程组，只是这时要想法消去两个未知量，才能得到一元一次方程.

（2）方程组解的状况的讨论. 在平面解析几何中，方程组(1)的一组解代表两条直线的一个公共点. 因此，不难看出，方程组(1)

的解可能出现下列三种情况：(i) 两直线相交，这时方程组有唯一解；(ii) 两直线平行而不重合，这时方程组无解；(iii) 两直线重合，这时方程组有无穷多组解. 同样，因为在空间直角坐标系内，每个三元线性方程代表一个平面，所以，给定如下三元线性方程组

$$\begin{cases} a_1x + b_1y + c_1z = d_1, \\ a_2x + b_2y + c_2z = d_2, \\ a_3x + b_3y + c_3z = d_3, \end{cases} \tag{2}$$

则它的每一组解代表三个平面的一个公共点. 利用几何直观不难看出，方程组(2)的解可能出现下列三种情况：(i) 三个平面相交于一点(例如三个坐标平面相交于坐标原点)，这时方程组有唯一解；(ii) 若三个平面中有某两个平行而不重合，这时方程组无解；(iii) 三个平面相交于一直线或三个平面互相重合，或两个平面重合，与第三个平面交于一条直线，这时方程组有无穷多组解.

在自然科学和工程技术中，常常需要处理几十、几百甚至成千上万个未知量的线性方程组，这就要求我们把关于二元和三元线性方程组的上述两个方面的知识推广到有 n 个未知量的线性方程组上去. 在这一章里，我们就要求解决这个问题. 具体地说，我们将要对 n 个未知量的线性方程组做如下两个方面的研究：

第一方面，是从理论上探讨下列三个问题：(1) 一个线性方程组在什么情况下有解，什么情况下无解？(2) 若有解，则有多少组解？(3) 在有许多组解(例如有无穷多组解)的情况下，解与解之间存在什么关系？

第二方面，是对有解的线性方程组探讨求解的方法，也就是把上面提到的消元法理论化、规格化，使它适用于有许多未知量的线性方程组.

由于上述第二方面的问题比较简单，所以，我们将首先探讨第二方面的问题. 即先研究线性方程组的求解方法，然后再回过头来研究线性方程组的一般理论问题.

§1 矩阵消元法

现在我们来介绍求解线性方程组的矩阵消元法. 这个方法是一个古典的方法,具有悠久的历史,但由于它行之有效,至今仍然是求解线性方程组的最基本方法之一. 这个方法中所包含的基本思想,在线性代数的其它一系列理论问题和计算问题中也将发挥重要的作用.

因为我们下面将要研究的线性方程组具有 n 个未知量,我们不可能用 $x,y,z,\cdots$ 等不同字母来代表它. 因此,下面我们用一个字母带上不同下角标来代表不同的未知量. 例如,用 $x_1,x_2,\cdots,x_n$ 代表 n 个未知量. 这时,有 n 个未知量 m 个方程的线性方程组的一般形式可表示为

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases} \tag{1}$$

在这个方程组中,未知量前面的系数带有两个下角标,第一个下角标代表它在第几个方程,第二个下角标代表它是第几个未知量的系数. 所以,a_{ij}代表的是第 i 个方程中未知量 x_j 的系数. 方程组(1)中等式右端的 $b_1,b_2,\cdots,b_m$ 称为**常数项**. 在这个方程组中,方程的个数 m 没有限制,可以小于 n,可以等于 n,也可以大于 n.

在线性方程组(1)中,如果让未知量取一组确定的数值:

$$x_1 = k_1,\ x_2 = k_2,\ \cdots,\ x_n = k_n,$$

代入方程组后使它转化为恒等式,则这一组数称为方程组(1)的一组**解**.

下面,我们就来具体讨论线性方程组(1)的求解方法.

矩阵消元法的原理

所谓矩阵消元法，是以下面的定义和一个命题作为理论基础的一种解线性方程组(1)的方法.

定义 方程组(1)做如下三种变换：

(i) 互换两个方程的位置；

(ii) 把某一个方程两边同乘一个非零常数 c；

(iii) 把某一个方程加上另一方程的 k 倍.

上述三种变换中的每一种都称为线性方程组(1)的**初等变换**.

应当指出：线性方程组的初等变换是可逆的. 也就是说，如果经过一次初等变换把方程组(1)变成一个新方程组，那么，新方程组必可经一次初等变换变为原方程组(1). 这可以具体讨论如下：

1. 如果互换方程组(1)中第 i,j 两方程的位置，则对新方程组再互换 i,j 两方程的位置就变回原方程组(1)；

2. 如果把方程组(1)的第 i 个方程乘以非零常数 c，那么，只要把新方程组的第 i 个方程乘以 $1/c$ 就变回原方程组(1)；

3. 如果把方程组(1)的第 j 个方程加上第 i 个方程的 k 倍，那么，只要把新方程组的第 j 个方程加上第 i 个方程的 $-k$ 倍就变回原方程组(1).

显然，如果方程组(1)经过若干次初等变换化为一个新方程组，那么，新方程组也可以经若干次初等变换化为原方程组(1). 另外还应指出，在做初等变换的过程中，方程组中方程的个数既不增加，也不减少.

命题 1.1 设方程组(1)经过某一初等变换后变为另一个方程组，则新方程组与原方程组同解.

证 设方程组(1)有一组解

$$x_1 = k_1,\ x_2 = k_2,\ \cdots,\ x_n = k_n, \tag{2}$$

代入(1)之后得到 m 个恒等式

$$\begin{cases} a_{11}k_1 + a_{12}k_2 + \cdots + a_{1n}k_n = b_1, \\ a_{21}k_1 + a_{22}k_2 + \cdots + a_{2n}k_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}k_1 + a_{m2}k_2 + \cdots + a_{mn}k_n = b_m. \end{cases}$$

对这组恒等式做相同的变换，得到一组新恒等式（也有 m 个），它恰好是把(2)式代入新方程组所得的结果. 由此可知：原方程组(1)的任一组解都是新方程组的解.

反过来，因为新方程组也可以经过适当的初等变换化为原方程组(1)，所以按同样的道理，新方程组的任一组解也是原方程组(1)的解. 于是两方程组同解. ▌

下面举几个例子来具体说明如何利用命题 1.1 来解线性方程组.

例 1 解线性方程组

$$\begin{cases} 2x_2 - x_3 = 1, \\ x_1 - x_2 + x_3 = 0, \\ 2x_1 + x_2 - x_3 = -2. \end{cases}$$

解 对方程组相继做初等变换.

(i) 调换第 1,2 两方程的位置，得

$$\begin{cases} x_1 - x_2 + x_3 = 0, \\ 2x_2 - x_3 = 1, \\ 2x_1 + x_2 - x_3 = -2. \end{cases}$$

(ii) 把上面的方程组第 3 个方程加上第 1 个方程的 -2倍，得

$$\begin{cases} x_1 - x_2 + x_3 = 0, \\ 2x_2 - x_3 = 1, \\ 3x_2 - 3x_3 = -2. \end{cases}$$

上面两次初等变换的目的是使方程组第 1 个方程保留 x_1，而第 2,3 个方程中未知量 x_1 都不出现（即其系数为零）.

(iii) 把(ii)中的方程组的第 3 个方程加上第 2 个方程的 -1

倍，得

$$\begin{cases} x_1 - x_2 + x_3 = 0, \\ 2x_2 - x_3 = 1, \\ x_2 - 2x_3 = -3. \end{cases}$$

(iv) 再把上面的方程组中第 2，3 两方程对换位置，得

$$\begin{cases} x_1 - x_2 + x_3 = 0, \\ x_2 - 2x_3 = -3, \\ 2x_2 - x_3 = 1. \end{cases}$$

(v) 把上面的方程组的第 3 个方程加上第 2 个方程的 −2 倍，得

$$\begin{cases} x_1 - x_2 + x_3 = 0, \\ x_2 - 2x_3 = -3, \\ 3x_3 = 7. \end{cases}$$

最后所得的方程组具有这样的特点：自上而下看，未知量的个数依次减少，成为阶梯形状(上面用虚线标出阶梯形). 只要从它的第 3 个方程解出 x_3，代入第 2 个方程解出 x_2，再代入第 1 个方程解出 x_1，就得到方程组的解

$$x_1 = -\frac{2}{3}, \quad x_2 = \frac{5}{3}, \quad x_3 = \frac{7}{3}.$$

根据命题 1.1，上面各个步骤中所得的每一个方程组都与原方程组同解，故最后方程组的这一组唯一的解就是原方程组的唯一解.

上面这个简单例子代表了用消元法解线性方程组的一般方法和计算格式. 它的基本思想是：反复利用方程组的初等变换以把方程组转化成阶梯形状的方程组.

读者不难发觉到，在上面的运算过程中，只是对方程组的系数进行运算，而未知量 x_1, x_2, x_3 在整个过程中未参加任何计算. 因此，每一步都把它们逐一写出完全是多余的累赘. 在计算中完全可以把它们先隐去. 只是这时要注意不要打乱了系数的排列顺序. 基于这一认识，我们把例 1 的方程组简化成如下的 3 行 4 列长方表

格

$$\begin{pmatrix} 0 & 2 & -1 & 1 \\ 1 & -1 & 1 & 0 \\ 2 & 1 & -1 & -2 \end{pmatrix}.$$

这个表格称为一个 3 行 4 列矩阵，简称为 3×4 矩阵. 它的每个横排称为行，竖排称为列. 现在，它的每一行代表原方程组的一个方程，第 1，2，3 列分别代表各方程中 x_1, x_2, x_3 的系数，第 4 列代表常数项.

于是，例 1 的解方程的各个步骤现在可简写成如下形式（用箭头⟶表示一次初等变换）

$$\begin{pmatrix} 0 & 2 & -1 & 1 \\ 1 & -1 & 1 & 0 \\ 2 & 1 & -1 & -2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & 2 & -1 & 1 \\ 2 & 1 & -1 & 2 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & 2 & -1 & 1 \\ 0 & 3 & -3 & -2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & 2 & -1 & 1 \\ 0 & 1 & -2 & -3 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & 1 & -2 & -3 \\ 0 & 2 & -1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 3 & 7 \end{pmatrix}.$$

等到把矩阵变成阶梯形后，再写出它代表的方程组

$$\begin{cases} x_1 - x_2 + x_3 = 0, \\ \qquad x_2 - 2x_3 = -3, \\ \qquad\qquad 3x_3 = 7. \end{cases}$$

求解最后的阶梯形方程组即得原方程组的全部解. 这种方法就称为**矩阵消元法**.

在熟悉了矩阵消元法的思路和计算格式之后，在实际计算时可把几次初等变换一步完成.

例 2 解方程组

$$\begin{cases} x_1 + x_2 - 3x_3 = -1, \\ 2x_1 + x_2 - 2x_3 = 1, \\ x_1 + x_2 + x_3 = 3, \\ x_1 + 2x_2 - 3x_3 = 3. \end{cases}$$

解 利用矩阵消元法

$$\begin{pmatrix} 1 & 1 & -3 & -1 \\ 2 & 1 & -2 & 1 \\ 1 & 1 & 1 & 3 \\ 1 & 2 & -3 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & -3 & -1 \\ 0 & -1 & 4 & 3 \\ 0 & 0 & 4 & 4 \\ 0 & 1 & 0 & 4 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 1 & -3 & -1 \\ 0 & 1 & -4 & -3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 4 & 7 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & -3 & -1 \\ 0 & 1 & -4 & -3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

其中第 1 个箭头表示利用第 1 行消去第 2,3,4 行中 x_1 的系数;第 2 个箭头表示:把第 2 行加到第 4 行,然后以(−1)乘第 2 行,最后再以(1/4)乘第 3 行,等等.在计算中要注意初等变换的先后次序,以免混乱,出现错误.例如,在上面第 2 个箭头的计算中,把第 2 行加到第 4 行后,第 4 行已经变成(0 0 4 7),只能以它为基础接下去再做变换.如果忽略这一点,还用原来的第 4 行又加到第 2 行,把整个矩阵变成

$$\begin{pmatrix} 1 & 1 & -3 & -1 \\ 0 & 0 & 4 & 7 \\ 0 & 0 & 4 & 4 \\ 0 & 0 & 4 & 7 \end{pmatrix},$$

这就错了.

现在把最后阶梯形矩阵对应的方程组写出

$$\begin{cases} x_1 + x_2 - 3x_3 = -1, \\ x_2 - 4x_3 = -3, \\ x_3 = 1, \\ 0 = 3. \end{cases}$$

这是一个矛盾方程组，无解. 故原方程组也无解.

例 3 讨论下面线性方程组解的状况

$$\begin{cases} x_1 + x_2 - 3x_4 - x_5 = 0, \\ x_1 - x_2 + 2x_3 - x_4 = 0, \\ 4x_1 - 2x_2 + 6x_3 + 3x_4 - 4x_5 = 0, \\ 2x_1 + 4x_2 - 2x_3 + 4x_4 - 7x_5 = 0. \end{cases}$$

解 这个方程组的常数项都是零，在消元过程中显然也永远是零，因此可以不写出来，只要写出未知量的系数所成的矩阵就可以了

$$\begin{pmatrix} 1 & 1 & 0 & -3 & -1 \\ 1 & -1 & 2 & -1 & 0 \\ 4 & -2 & 6 & 3 & -4 \\ 2 & 4 & -2 & 4 & -7 \end{pmatrix}.$$

现在有 5 个未知量，计算显然复杂一些. 为了避免计算中出错，可以采用加校正项的办法. 这就是：把上面矩阵中每一行元素加在一起放在该行的最后，得到一个 4 行 6 列的新矩阵(为防止校正项和前面的方程组系数混淆，可用虚线隔开)，然后对新矩阵做消元法. 在消元过程中，每一步都保持着每行前 4 个数之和等于第 5 个数的性质. 如发现某一行破坏了这一点，那就证明该行计算错了(这个方法对常数项不为零的方程组也适用).

$$\left(\begin{array}{ccccc:c} 1 & 1 & 0 & -3 & -1 & -2 \\ 1 & -1 & 2 & -1 & 0 & 1 \\ 4 & -2 & 6 & 3 & -4 & 7 \\ 2 & 4 & -2 & 4 & -7 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{rrrrr:r} 1 & 1 & 0 & -3 & -1 & -2 \\ 0 & -2 & 2 & 2 & 1 & 3 \\ 0 & -6 & 6 & 15 & 0 & 15 \\ 0 & 2 & -2 & 10 & -5 & 5 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{rrrrr:r} 1 & 1 & 0 & -3 & -1 & -2 \\ 0 & -2 & 2 & 2 & 1 & 3 \\ 0 & 0 & 0 & 9 & -3 & 6 \\ 0 & 0 & 0 & 12 & -4 & 8 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{rrrrr:r} 1 & 1 & 0 & -3 & -1 & -2 \\ 0 & 2 & -2 & -2 & -1 & -3 \\ 0 & 0 & 0 & 3 & -1 & 2 \\ 0 & 0 & 0 & 3 & -1 & 2 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{rrrrr:r} 1 & 1 & 0 & -3 & -1 & -2 \\ 0 & 2 & -2 & -2 & -1 & -3 \\ 0 & 0 & 0 & 3 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

再写出它所对应的方程组(注意常数项为零)

$$\begin{cases} x_1 + x_2 \qquad - 3x_4 - x_5 = 0, \\ \qquad 2x_2 - 2x_3 - 2x_4 - x_5 = 0, \\ \qquad\qquad\qquad 3x_4 - x_5 = 0. \end{cases}$$

最后一个方程是 $0=0$,可以不写出来.现在把位于阶梯形的角上的未知量 x_1,x_2,x_4 保留在方程的左端,其余未知量 x_3,x_5 移到方程的右端,得

$$\begin{cases} x_1 + x_2 - 3x_4 = \qquad x_5, \\ \qquad 2x_2 - 2x_4 = 2x_3 + x_5, \\ \qquad\qquad 3x_4 = \qquad x_5. \end{cases}$$

在这个方程组中,x_3 与 x_5 任取一组值代入,都能唯一地确定出 x_1,x_2,x_4 的值,从而得到原方程组的一组解. x_3,x_5 称为**自由未知**

量. 这时方程组有无穷多组解.

在用矩阵消元法解线性方程组时，我们实际上可以做的更彻底一些，就是把得到的阶梯形矩阵每一行左边第一个不为零的数化成 1，再用它消去它顶上各行相应的元素. 例如对例 3 最后的阶梯形矩阵再做此消元：

$$\left(\begin{array}{ccccc:c} 1 & 1 & 0 & -3 & -1 & -2 \\ 0 & 2 & -2 & -2 & -1 & -3 \\ 0 & 0 & 0 & 3 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{ccccc:c} 1 & 1 & 0 & -3 & -1 & -2 \\ 0 & 2 & -2 & -2 & -1 & -3 \\ 0 & 0 & 0 & 1 & -1/3 & 2/3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccccc:c} 1 & 1 & 0 & 0 & -2 & 0 \\ 0 & 2 & -2 & 0 & -5/3 & -5/3 \\ 0 & 0 & 0 & 1 & -1/3 & 2/3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{ccccc:c} 1 & 1 & 0 & 0 & -2 & 0 \\ 0 & 1 & -1 & 0 & -5/6 & -5/6 \\ 0 & 0 & 0 & 1 & -1/3 & 2/3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccccc:c} 1 & 0 & 1 & 0 & -7/6 & 5/6 \\ 0 & 1 & -1 & 0 & -5/6 & -5/6 \\ 0 & 0 & 0 & 1 & -1/3 & 2/3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

与最后阶梯形矩阵相应的方程组是

$$\begin{cases} x_1 + x_3 - 7/6x_5 = 0, \\ x_2 - x_3 - 5/6x_5 = 0, \\ x_4 - 1/3x_5 = 0. \end{cases}$$

由此立得原方程组解的一般表达式：

$$x_1 = -x_3 + \frac{7}{6}x_5, \quad x_2 = x_3 + \frac{5}{6}x_5, \quad x_4 = \frac{1}{3}x_5.$$

在例 3 的消元过程中第 2 个箭头处发生了一个跳跃现象，即当我们利用第 2 行消去第 3,4 行中 x_2 的系数时，第 3,4 行中 x_3 的系数也都变成零. 这种跳跃现象实际上可以消除. 这只要在原方程组中把未知量重新编号就可以了. 具体地说，在例 3 中只要把 x_3 与 x_4 互换位置，即设 $y_1=x_1$，$y_2=x_2$，$y_3=x_4$，$y_4=x_3$，$y_5=x_5$，

方程组变成

$$\begin{cases} y_1 + y_2 - 3y_3 - y_5 = 0, \\ y_1 - y_2 - y_3 + 2y_4 = 0, \\ 4y_1 - 2y_2 + 3y_3 + 6y_4 - 4y_5 = 0, \\ 2y_1 + 4y_2 + 4y_3 - 2y_4 - 7y_5 = 0. \end{cases}$$

再做消元，最后得到的阶梯形矩阵是

$$\begin{pmatrix} 1 & 1 & -3 & 0 & -1 \\ 0 & 2 & -2 & -2 & -1 \\ 0 & 0 & 3 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

跳跃的现象消除了，出现了规则的阶梯形. 它对应的方程组是

$$\begin{cases} y_1 + y_2 - 3y_3 - y_5 = 0, \\ 2y_2 - 2y_3 - 2y_4 - y_5 = 0, \\ 3y_3 - y_5 = 0. \end{cases}$$

移项，得

$$\begin{cases} y_1 + y_2 - 3y_3 = y_5, \\ 2y_2 - 2y_3 = 2y_4 + y_5, \\ 3y_3 = y_5. \end{cases}$$

现在 y_4 与 y_5 是自由未知量.

线性方程组解的初步讨论

现在利用矩阵消元法来对线性方程组做理论上的初步讨论. 首先给出矩阵的一般定义.

定义 给定 mn 个数 $a_{ij}(i=1,2,\cdots,m;j=1,2,\cdots,n)$，把它们按一定次序排成一个 m 行 n 列的长方形表格

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

称为一个 m **行** n **列矩阵**,简称为 $m\times n$ **矩阵**.

在一个矩阵中,横向的一排数称为矩阵的**行**,竖向的一排数称为矩阵的**列**. 矩阵中的每个数称为它的**元素**. 在上面的矩阵中,每个元素 a_{ij}带有两个下角标,第 1 个下角标代表它所在的行,第 2 个下角标代表它所在的列. a_{ij}表示该元素位于矩阵的第 i 行和第 j 列的交叉点处.

矩阵是一个独立的概念,并不一定要跟线性方程组联系在一起. 但是它是研究线性方程组的有力工具. 现在矩阵对我们来说还仅仅看成是一个"表格",在下面,我们将逐步赋于它新的内容,把它丰富起来,从而使它发挥越来越大的作用.

有了上面的定义之后,方程组(1)中未知量的系数就可以排成一个矩阵 A,它就是定义中所写出的那个 $m\times n$ 矩阵,我们称 A 为方程组(1)的**系数矩阵**. 如果把(1)的常数项添到 A 内作为最后一列,就得到一个 $m\times(n+1)$矩阵

$$\overline{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix},$$

矩阵 $\overline{A}$ 称为方程组(1)的**增广矩阵**.

我们不妨假定 $\overline{A}$ 的前 n 列中任一列元素都不全为零(因为如果第 j 列元素 $a_{ij}(i=1,2,\cdots,m)$全为零,那么,未知量 x_j 在方程组中实际上不出现,这情况可预先排除在外). 现在对 $\overline{A}$ 做矩阵消元法:

(i) 通过调换两行的位置以保持第 1 行第 1 列处元素不为零. 因为第 1 列元素不全为零,所以总能做到这一点. 例 1 的第 1 步初等变换就是例子.

(ii) $a_{11}\neq 0$ 时,利用第 1 行乘以适当倍数加到第 2,3,$\cdots$,m 行,把矩阵变成

$$\overline{A} \longrightarrow \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & b_{22} & \cdots & b_{2n} & b_2' \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & b_{m2} & \cdots & b_{mn} & b_m' \end{pmatrix}.$$

下面，再对用虚线标出的右下角部分继续上述变换. 这时可能出现例 3 所指出的跳跃现象，比如出现

$$b_{22} = b_{32} = \cdots = b_{m2} = 0$$

的情况. 但在例 3 中已指出，只要重新对未知量进行编号就可以消除这种现象. 我们可以设想像例 3 那样，先把 $\overline{A}$ 化成阶梯形，再重排一下未知量的号码以达到这一目的. 为了讨论方便，不妨假定消元过程中不出现跳跃. 于是不断重复(i)，(ii)两个步骤，最后把 $\overline{A}$ 变为如下阶梯形矩阵

$$\overline{A} \longrightarrow \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1r} & \cdots & c_{1n} & d_1 \\ 0 & c_{22} & \cdots & c_{2r} & \cdots & c_{2n} & d_2 \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & c_{rr} & \cdots & c_{rn} & d_r \\ 0 & 0 & \cdots & 0 & \cdots & 0 & d_{r+1} \\ 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \end{pmatrix},$$

其中 $c_{11}, c_{22}, \cdots, c_{rr}$ 均不为零. 写出它所对应的方程组

$$\begin{cases} c_{11}x_1 + c_{12}x_2 + \cdots + c_{1r}x_r + \cdots + c_{1n}x_n = d_1, \\ \qquad\quad c_{22}x_2 + \cdots + c_{2r}x_r + \cdots + c_{2n}x_n = d_2, \\ \qquad\qquad\qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \qquad\qquad\qquad\qquad c_{rr}x_r + \cdots + c_{rn}x_n = d_r, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad 0 = d_{r+1}. \end{cases}$$

(注意最后阶梯形矩阵的前 n 列正好是由方程组(1)的系数矩阵 A

变化而来的.)

现在可以得到如下结论:

1. 如果 $d_{r+1}\neq 0$,即系数矩阵 A 变化后的阶梯的个数 r 和增广矩阵 $\overline{A}$ 变化后的阶梯个数 $r+1$ 不相等,则方程组(1)无解.前面例 2 就属这种情况.

2. 如果 $d_{r+1}=0$,即 A 与 $\overline{A}$ 变化后所得阶梯数都是 r,则方程组有解.

3. 在 $d_{r+1}=0$,方程组有解时,如果 $r=n$,那么最后得到的方程组是

$$\begin{cases} c_{11}x_1 + c_{12}x_2 + \cdots + c_{1n}x_n = d_1, \\ \qquad\quad c_{22}x_2 + \cdots + c_{2n}x_n = d_2, \\ \qquad\qquad\qquad \cdots\cdots\cdots\cdots \\ \qquad\qquad\qquad\quad c_{nn}x_n = d_n. \end{cases}$$

自下而上依次确定出 $x_n, x_{n-1}, \cdots, x_2, x_1$ 的值,此时方程组显然只有唯一的一组解.前面例 1 属此情况.

4. 在 $d_{r+1}=0$,且 $r<n$ 时,所得方程组为

$$\begin{cases} c_{11}x_1 + c_{12}x_2 + \cdots + c_{1r}x_r + c_{1\,r+1}x_{r+1} + \cdots + c_{1n}x_n = d_1, \\ \qquad\quad c_{22}x_2 + \cdots + c_{2r}x_r + c_{2\,r+1}x_{r+1} + \cdots + c_{2n}x_n = d_2, \\ \qquad\qquad\qquad \cdots\cdots\cdots\cdots\cdots\cdots \\ \qquad\qquad\qquad\quad c_{rr}x_r + c_{r\,r+1}x_{r+1} + \cdots + c_{rn}x_n = d_r. \end{cases}$$

移项,得

$$\begin{cases} c_{11}x_1 + c_{12}x_2 + \cdots + c_{1r}x_r = d_1 - (c_{1\,r+1}x_{r+1} + \cdots + c_{1n}x_n), \\ \qquad\quad c_{22}x_2 + \cdots + c_{2r}x_r = d_2 - (c_{2\,r+1}x_{r+1} + \cdots + c_{2n}x_n), \\ \qquad\qquad\qquad \cdots\cdots\cdots\cdots\cdots\cdots \\ \qquad\qquad\qquad\quad c_{rr}x_r = d_r - (c_{r\,r+1}x_{r+1} + \cdots + c_{rn}x_n). \end{cases}$$

任给 $x_{r+1}, \cdots, x_n$ 一组值,代入上方程组都可唯一确定出 $x_r, x_{r-1}, \cdots, x_1$ 的一组值.这时方程组有无穷多组解,而 $x_{r+1}, \cdots, x_n$ 称为**自由未知量**,其值可任取.自由未知量个数为 $n-r$.前面例 3 属此情

况. 在这种情况下，可把解表成自由未知量 $x_{r+1},\cdots,x_n$ 的函数，这种解的表达式称为方程组的**一般解**. 前面例 3 所得到的解的表达式就是该方程组的一般解.

综合以上的讨论，可得如下结论：

1. $d_{r+1}\neq 0$ 时方程组无解；
2. $d_{r+1}=0$，且 $r=n$ 时方程组有唯一解；
3. $d_{r+1}=0$，且 $r<n$ 时方程组有无穷多组解.

例 4 解线性方程组

$$\begin{cases} x_1+x_2+x_3+x_4+x_5=7, \\ 3x_1+2x_2+x_3+x_4-3x_5=-2, \\ x_2+2x_3+2x_4+6x_5=23, \\ 5x_1+4x_2+3x_3+3x_4-x_5=12. \end{cases}$$

解 用矩阵消元法.

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 7 & \vdots & 12 \\ 3 & 2 & 1 & 1 & -3 & -2 & \vdots & 2 \\ 0 & 1 & 2 & 2 & 6 & 23 & \vdots & 34 \\ 5 & 4 & 3 & 3 & -1 & 12 & \vdots & 26 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 7 & \vdots & 12 \\ 0 & -1 & -2 & -2 & -6 & -23 & \vdots & -34 \\ 0 & 1 & 2 & 2 & 6 & 23 & \vdots & 34 \\ 0 & -1 & -2 & -2 & -6 & -23 & \vdots & -34 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 7 & \vdots & 12 \\ 0 & 1 & 2 & 2 & 6 & 23 & \vdots & 34 \\ 0 & 0 & 0 & 0 & 0 & 0 & \vdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \vdots & 0 \end{pmatrix}.$$

写出对应方程组

$$\begin{cases} x_1+x_2+x_3+x_4+x_5=7, \\ x_2+2x_3+2x_4+6x_5=23. \end{cases}$$

移项，得

$$\begin{cases} x_1 + x_2 = 7 - (x_3 + x_4 + x_5), \\ \qquad x_2 = 23 - (2x_3 + 2x_4 + 6x_5). \end{cases}$$

故得方程组的一般解

$$\begin{cases} x_1 = -16 + x_3 + x_4 + 5x_5, \\ x_2 = 23 - 2x_3 - 2x_4 - 6x_5. \end{cases}$$

其中 x_3, x_4, x_5 为自由未知量，其值可任取.

齐次线性方程组

如果一个线性方程组的常数项都是零，就称为**齐次线性方程组**. 例 3 就是一个齐次线性方程组. 齐次线性方程组的一般形式是

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0. \end{cases} \tag{3}$$

方程组(3)显然有一组解

$$x_1 = 0, \quad x_2 = 0, \quad \cdots, \quad x_n = 0.$$

这组解称为**零解**或**平凡解**. 除此之外的其它解(如果存在的话)称为**非零解**或**非平凡解**.

由此可知:齐次线性方程组总是有解的. 这从前面对方程组的一般性讨论中也不难看出. 因为齐次线性方程组常数项全为零，消元后所得的阶梯形矩阵中，$d_1 = d_2 = \cdots = d_r = d_{r+1} = 0$，所以永远属于有解的情况.

从上面对方程组的一般性讨论中可以立刻得到关于齐次线性方程组的如下结论：

1. 如消元后阶梯形矩阵中 $r = n$，则齐次线性方程组只有唯一的一组解，因而只有零解；

2. 如果 $r < n$，则齐次线性方程组有无穷多组解，从而有非零解.

因为 $r \leqslant m$，所以，如果 $m < n$，则必 $r < n$，于是有：

命题 1.2 如果齐次线性方程组(3)的方程个数 m 小于未知量个数 n,则它必有非零解.

例如例 3 中的齐次线性方程组有 4 个方程 5 个未知量,不必求解就可以判断它必定有非零解. 这与实际求解所得的结果是一致的. 命题 1.2 将在下一节中使用.

在本节的最后,我们提请读者注意一个有趣的现象:在利用矩阵消元法解线性方程组时,只进行加、减、乘、除四则运算,没有涉及开方运算,更没有涉及解二次以上代数方程的计算. 因此:

1. 如果线性方程组(1)的系数 a_{ij}和常数项 $b_i(i=1,2,\cdots,m;j=1,2,\cdots,n)$都是有理数,那么计算过程中就不会出现无理数. 如果限定方程组一般解中的自由未知量(如果有的话)也只取有理数的值,那么,我们得到的解也全是有理数. 从而,我们可以把对这样的方程组的讨论限制在有理数的范围内. 其所以如此,是因为有理数经过加、减、乘、除四则运算后仍然是有理数,我们称有理数对加、减、乘、除运算是**封闭的**.

2. 如果线性方程组(1)的系数和常数项全是实数,因为实数对加、减、乘、除四则运算也是封闭的,所以,对这样的方程组的讨论也可以限制在实数的范围内.

上述现象在代数学中有重要的意义. 我们在本章的最后再来对它做进一步的阐述.

习 题 一

1. 用矩阵消元法求下列线性方程组的解:

(1) $$\begin{cases} 2x_1-x_2+x_3-x_4=1, \\ 2x_1-x_2-3x_4=2, \\ 3x_1-x_3+x_4=-3, \\ 2x_1+2x_2-2x_3+5x_4=-6. \end{cases}$$

(2) $\begin{cases} x_1+3x_2+5x_3-4x_4 = 1, \\ x_1+3x_2+2x_3-2x_4+x_5=-1, \\ x_1-2x_2+x_3-x_4-x_5=3, \\ x_1-4x_2+x_3+x_4-x_5=3, \\ x_1+2x_2+x_3-x_4+x_5=-1. \end{cases}$

(3) $\begin{cases} x_1+2x_2-3x_4+2x_5=1, \\ x_1-x_2-3x_3+x_4-3x_5=2, \\ 2x_1-3x_2+4x_3-5x_4+2x_5=7, \\ 9x_1-9x_2+6x_3-16x_4+2x_5=25. \end{cases}$

(4) $\begin{cases} x_1-2x_2+3x_3-4x_4=4, \\ x_2-x_3+x_4=-3, \\ x_1+3x_2+x_4=1, \\ -7x_2+3x_3+x_4=-3. \end{cases}$

(5) $\begin{cases} 3x_1+4x_2-5x_3+7x_4=0, \\ 2x_1-3x_2+3x_3-2x_4=0, \\ 4x_1+11x_2-13x_3+16x_4=0, \\ 7x_1-2x_2+x_3+3x_4=0. \end{cases}$

(6) $\begin{cases} 2x_1+x_2-x_3+x_4=1, \\ 3x_1-2x_2+2x_3-3x_4=2, \\ 5x_1+x_2-x_3+2x_4=-1, \\ 2x_1-x_2+x_3-3x_4=4. \end{cases}$

(7) $\begin{cases} x_1+2x_2+3x_3-x_4=1, \\ 3x_1+2x_2+x_3-x_4=1, \\ 2x_1+3x_2+x_3+x_4=1, \\ 2x_1+2x_2+2x_3-x_4=1, \\ 5x_1+5x_2+2x_3=2. \end{cases}$

2. 证明齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = 0, \\ a_{21}x_1 + a_{22}x_2 = 0 \end{cases}$$

有非零解的充分必要条件是

$$a_{11}a_{22} - a_{12}a_{21} = 0.$$

3. 判断下列齐次线性方程组是否有非零解：

(1) $\begin{cases} 2x_1+3x_2- x_3+5x_4=0, \\ 3x_1- x_2+2x_3-7x_4=0, \\ 4x_1+ x_2-3x_3+6x_4=0, \\ x_1-2x_2+4x_3-7x_4=0. \end{cases}$

(2) $\begin{cases} x_2- x_3+ x_4=0, \\ -7x_2+3x_3+ x_4=0, \\ x_1+3x_2 -3x_4=0, \\ x_1-2x_2+3x_3-4x_4=0. \end{cases}$

(3) $\begin{cases} x_1-2x_2+x_3- x_4+ x_5=0, \\ 2x_1+ x_2-x_3+2x_4-3x_5=0, \\ 3x_1-2x_2-x_3+ x_4-2x_5=0, \\ 2x_1-5x_2+x_3-2x_4+2x_5=0. \end{cases}$

(4) $\begin{cases} x_1-x_2 =0, \\ x_2-x_3 =0, \\ x_3-x_4=0, \\ -x_1 +x_4=0. \end{cases}$

4. 证明空间中三向量

$$\alpha_1 = (a_{11},a_{21},a_{31}), \quad \alpha_2 = (a_{12},a_{22},a_{32}),$$
$$\alpha_3 = (a_{13},a_{23},a_{33})$$

共面的充分必要条件是齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = 0, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = 0, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = 0 \end{cases}$$

有非零解.

5. 讨论 λ,a,b 取何值时下列方程组有解，在有解时求其解.

(1) $\begin{cases}\lambda x_1+x_2+x_3=1,\\ x_1+\lambda x_2+x_3=\lambda,\\ x_1+x_2+\lambda x_3=\lambda^2.\end{cases}$ (2) $\begin{cases}ax_1+x_2+x_3=4,\\ x_1+bx_2+x_3=3,\\ x_1+2bx_2+x_3=4.\end{cases}$

(3) $\begin{cases}x_1+x_2+x_3+x_4+x_5=1,\\ 3x_1+2x_2+x_3+x_4-3x_5=a,\\ x_2+2x_3+2x_4+6x_5=3,\\ 5x_1+4x_2+3x_3+3x_4-x_5=b.\end{cases}$

6. 设

$$x_1-x_2=a_1,\quad x_2-x_3=a_2,\quad x_3-x_4=a_3,$$
$$x_4-x_5=a_4,\quad x_5-x_1=a_5.$$

证明:这方程组有解的充分必要条件是

$$\sum_{i=1}^{5}a_i=0.$$

在有解的情况下,求其一般解.

§2 n维向量空间

为了从理论上深入地讨论线性方程组解的性质,我们必须对线性方程组本身做进一步的分析,以便更深刻地认识它的本质.

我们从分析三个方程的线性方程组入手.设给定如下实系数线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1,\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2,\\ a_{31}x_1+a_{32}x_2+\cdots+a_{3n}x_n=b_3.\end{cases}$$

回想在空间解析几何中,每一个向量在空间直角坐标系下可表示为一个三元有序数组

$$\alpha=(a_1,a_2,a_3).$$

为了讨论问题方便,我们也常常把它竖起来写

$$\alpha = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}.$$

向量的加法和数乘的运算法则是

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \end{pmatrix};\ k\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} ka_1 \\ ka_2 \\ ka_3 \end{pmatrix}.$$

现在把上面的线性方程组的系数和常数项分别看作三维空间中的向量

$$\alpha_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix},\ \alpha_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix},\ \cdots,\ \alpha_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ a_{3n} \end{pmatrix},\ \beta = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

这实际也是把该方程组的系数矩阵 A 和增广矩阵 $\overline{A}$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \end{pmatrix},\ \overline{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ a_{31} & a_{32} & \cdots & a_{3n} & b_3 \end{pmatrix}$$

的列看成三维空间中的向量 $\alpha_1,\alpha_2,\cdots,\alpha_n,\beta$. 利用向量的加法和数乘运算，我们所探讨的方程组现在可以改写成向量方程的形状

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \beta.$$

由此可知：

（1）上述线性方程组是否有解等价于向量 β 能否用向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 表示出来；

（2）方程组有多少组解的问题等价于 β 能有多少种方式用 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 表示出来.

这样一来，我们就把有三个方程的线性方程组和解析几何中的向量联系起来了. 这使我们对线性方程组的认识深入了一步，这具有重要的意义.

当讨论更多个方程所组成的线性方程组时，解析几何中的向量就不够用了，需要我们从理论上加以推广. 下面就来做这个工作.

n 维向量空间的基本概念

定义　一个 n 元有序数组

$$\alpha = (a_1, a_2, \cdots, a_n)$$

称为一个 n **维向量**. $a_1, a_2, \cdots, a_n$ 称为向量 α 的**分量**或**坐标**.

现在向量的分量 $a_1, a_2, \cdots, a_n$ 也可以是复数. 一个 n 维向量根据讨论问题的需要，也可以竖起来写

$$\alpha = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

定义　n 维向量的加法和与数的数乘运算规定为:

(i) 加法

$$(a_1, a_2, \cdots, a_n) + (b_1, b_2, \cdots, b_n) = (a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n);$$

(ii) 对任意数 k，数乘

$$k(a_1, a_2, \cdots, a_n) = (ka_1, ka_2, \cdots, ka_n).$$

上面的定义显然只是三维空间向量加法和数乘运算的简单推广，只是现在所涉及的数已经不限于实数，而是包括复数在内. 如果向量改用竖的写法，那么加法和数乘就是

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix}; \quad k\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} ka_1 \\ ka_2 \\ \vdots \\ ka_n \end{pmatrix}.$$

两个 n 维向量相等是指它们的分量完全相同. 全体 n 维向量连同上面定义的加法和数乘运算合称一个 n **维向量空间**.

命题 2.1　n 维向量空间中向量的加法和数乘运算满足如下

八条规律：

(1) 加法结合律　$\alpha+(\beta+\gamma)=(\alpha+\beta)+\gamma$；

(2) 加法交换律　$\alpha+\beta=\beta+\alpha$；

(3) 称$(0,0,\cdots,0)$为 n 维**零向量**，记为 0. 对任一 n 维向量 α，有 $0+\alpha=\alpha+0=\alpha$；

(4) 任给 $\alpha=(a_1,a_2,\cdots,a_n)$，记 $-\alpha=(-a_1,-a_2,\cdots,-a_n)$，称其为 α 的**负向量**. 它满足 $\alpha+(-\alpha)=(-\alpha)+\alpha=0$；

(5) 对数 1，有 $1\cdot\alpha=\alpha$；

(6) 对任意数 k,l，有 $(kl)\alpha=k(l\alpha)$；

(7) 对任意数 k,l，有 $(k+l)\alpha=k\alpha+l\alpha$；

(8) 对任意数 k，有 $k(\alpha+\beta)=k\alpha+k\beta$.

其中希腊字母 α,β,γ 表示 n 维向量(下面均用希腊字母代表向量).

这个命题只需按定义逐一加以验证就可知其正确了. 我们这里把它列举出来，是要读者注意：这八条是 n 维向量空间的最基本的规律或性质. n 维向量空间的一系列基本命题都是以上述八条为基础推导出来的.

今后我们将 $\alpha+(-\beta)$写成 $\alpha-\beta$，称为向量的减法运算. 下面再介绍几个基本概念.

定义　给定一个 n 维向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$，又给定 s 个数 $k_1,k_2,\cdots,k_s$. 称向量 $k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s$ 为向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的一个**线性组合**.

定义　给定 n 维向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$. 设 β 是一个 n 维向量. 如果存在 s 个数 $k_1,k_2,\cdots,k_s$，使

$$\beta=k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s,$$

则称 β 可被向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ **线性表示**.

现在利用上面这个概念来分析一下 n 维向量空间中的向量和线性方程组之间的联系. 给定线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases} \tag{1}$$

考虑 m 维向量空间中的 $n+1$ 个向量

$$\alpha_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \ \alpha_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \ \cdots, \ \alpha_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}, \ \beta = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

应用 m 维向量的加法和数乘运算，方程组(1)可以改写成如下的向量方程

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \beta. \tag{2}$$

如果方程组(1)有一组解

$$x_1 = k_1, \ x_2 = k_2, \ \cdots, \ x_n = k_n,$$

代入(2)式，得

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n,$$

即 β 能被向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表示. 反之，若 β 能被向量组 α_1, $\alpha_2,\cdots,\alpha_n$ 线性表示，则表示的系数就是方程组(1)的一组解. 于是有如下两条结论：

1. 方程组(1)有解的充分必要条件是：向量 β 能被向量组 α_1, $\alpha_2,\cdots,\alpha_n$ 线性表示；

2. 方程组(1)的解的组数等于 β 被 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表示表法的种数.

上面的分析触及了线性方程组的本质. 读者应该十分熟悉线性方程组(1)和向量方程(2)之间的联系. 这就是说，给定一个线性方程组，读者应当立即能把它改写成向量方程. 反过来，给定一个向量方程，也应立即能把它改写成线性方程组的形式.

如果写出方程组(1)的系数矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

那么，$\alpha_1,\alpha_2,\cdots,\alpha_n$ 就是 A 中第 $1,2,\cdots,n$ 列. 因而，我们可以称 α_1，$\alpha_2,\cdots,\alpha_n$ 是矩阵 A 的列向量. 为方便计，可把 A 写成

$$A=(\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n),$$

它表示矩阵 A 是把 m 维向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 作为列向量依次排列而成的. 于是方程组(1)的增广矩阵可简记为

$$\overline{A}=(\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n \quad \beta).$$

另一方面，矩阵 A 的每个行则可以看作一个 n 维向量(用横排方式写出)，称为矩阵 A 的行向量. 这样的看法对讨论问题也是很有用的. 当然，矩阵的行向量和列向量一般是不同维数的向量.

例 1 将§1 例 4 的线性方程组

$$\begin{cases} x_1+x_2+x_3+x_4+x_5=7, \\ 3x_1+2x_2+x_3+x_4-3x_5=-2, \\ x_2+2x_3+2x_4+6x_5=23, \\ 5x_1+4x_2+3x_3+3x_4-x_5=12 \end{cases}$$

写成向量方程的形状. 为此，令

$$\alpha_1=\begin{pmatrix}1\\3\\0\\5\end{pmatrix},\quad \alpha_2=\begin{pmatrix}1\\2\\1\\4\end{pmatrix},\quad \alpha_3=\begin{pmatrix}1\\1\\2\\3\end{pmatrix},$$

$$\alpha_4=\begin{pmatrix}1\\1\\2\\3\end{pmatrix},\quad \alpha_5=\begin{pmatrix}1\\-3\\6\\-1\end{pmatrix},\quad \beta=\begin{pmatrix}7\\-2\\23\\12\end{pmatrix},$$

得向量方程

$$x_1\alpha_1+x_2\alpha_2+x_3\alpha_3+x_4\alpha_4+x_5\alpha_5=\beta.$$

如果在该例的一般解表达式中令 $x_3=x_4=x_5=1$，得到一组解

$$x_1=-9,\ x_2=13,\ x_3=1,\ x_4=1,\ x_5=1.$$

相应地得到 β 被 $\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5$ 线性表示的一种表示法

$$\beta=-9\alpha_1+13\alpha_2+\alpha_3+\alpha_4+\alpha_5.$$

这个方程组有无穷多组解，从而 β 有无穷多种方式表为 $\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5$ 的线性组合.

向量组的线性相关与线性无关

下面来介绍线性代数中两个最基本的概念：向量组的线性相关与线性无关概念.

定义 给定一个 m 维向量组

$$\alpha_1=\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{pmatrix},\quad \alpha_2=\begin{pmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\end{pmatrix},\ \cdots,\ \alpha_s=\begin{pmatrix}a_{1s}\\a_{2s}\\\vdots\\a_{ms}\end{pmatrix},$$

如果齐次线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1s}x_s=0,\\a_{21}x_1+a_{22}x_2+\cdots+a_{2s}x_s=0,\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\a_{m1}x_1+a_{m2}x_2+\cdots+a_{ms}x_s=0\end{cases}\tag{3}$$

有非零解，则称向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ **线性相关**；如果齐次线性方程组(3)只有零解，则称此向量组**线性无关**.

这就是说，如果我们以所给的向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 作为列向量排成一个矩阵

$$A=(\alpha_1\ \ \alpha_2\ \ \cdots\ \ \alpha_s)=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1s}\\a_{21}&a_{22}&\cdots&a_{2s}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{ms}\end{pmatrix},$$

那么，向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性相关或线性无关标志着，以 A 为系数矩阵的齐次线性方程组(3)是有非零解或没有非零解.

在 §1 中我们已经知道，可以用矩阵消元法来判断齐次线性方程组有无非零解，所以，现在就可以用它来判断一个向量组是否线性相关.

例 2 给定如下 4 维向量组

$$\alpha_1=(-1,\ 1,\ 2,\ 4),\qquad \alpha_2=(0,\ -1,\ 1,\ 3),$$
$$\alpha_3=(2,\ 0,\ 6,\ -2),\qquad \alpha_4=(-3,\ -1,\ 3,\ 4),$$
$$\alpha_5=(1,\ 1,\ -4,\ -7),$$

判断它们线性相关或线性无关.

解 把它们竖起来，排成一个 4×5 矩阵

$$A=\begin{pmatrix}-1 & 0 & 2 & -3 & 1\\ 1 & -1 & 0 & -1 & 1\\ 2 & 1 & 6 & 3 & -4\\ 4 & 3 & -2 & 4 & -7\end{pmatrix}.$$

以 A 为系数矩阵的齐次线性方程组为

$$\begin{cases}-x_1 \qquad\quad +2x_3-3x_4+x_5=0,\\ x_1-x_2 \qquad\quad -x_4+x_5=0,\\ 2x_1+x_2+6x_3+3x_4-4x_5=0,\\ 4x_1+3x_2-2x_3+4x_4-7x_5=0.\end{cases}$$

这个齐次线性方程组中方程个数为 4，小于未知量个数 5，由命题 1.2 知它必有非零解，从而知所给的向量组线性相关.

从这个例子可以归纳出更一般的结论：给定 s 个 m 维向量 α_1，$\alpha_2,\cdots,\alpha_s$. 如果 $s>m$，即向量个数大于向量的维数，那么，令

$$A=(\alpha_1\ \ \alpha_2\ \ \cdots\ \ \alpha_s),$$

以 A 为系数矩阵的齐次线性方程组有 m 个方程，s 个未知量，由命题 1.2 知它必有非零解，从而 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性相关.

例 3 给定向量组

$$\alpha_1=(7,0,0,0,0),\qquad \alpha_2=(-1,3,4,0,0),$$

$$\alpha_3=(1,0,1,1,0), \qquad \alpha_4=(0,0,1,1,-1),$$

判断它们是否线性相关.

解 把它们竖起来排成一个 5×4 矩阵

$$A=\begin{pmatrix}7&-1&1&0\\0&3&0&0\\0&4&1&1\\0&0&1&1\\0&0&0&-1\end{pmatrix}.$$

要判断以 A 为系数矩阵的齐次线性方程组是否有非零解，只需用矩阵消元法把 A 化为阶梯形

$$A\longrightarrow\begin{pmatrix}7&-1&1&0\\0&3&0&0\\0&0&1&1\\0&0&1&1\\0&0&0&-1\end{pmatrix}\longrightarrow\begin{pmatrix}7&-1&1&0\\0&1&0&0\\0&0&1&1\\0&0&0&0\\0&0&0&-1\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}7&-1&1&0\\0&1&0&0\\0&0&1&1\\0&0&0&1\\0&0&0&0\end{pmatrix}.$$

阶梯形矩阵中阶梯个数 $r=4$，未知量个数$=4$，因而相应的齐次线性方程组只有零解. 故 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 线性无关.

前面已经指出，线性方程组是否有解的问题等价于该方程组的常数向量是否能表示为其系数矩阵的列向量的线性组合的问题，所以向量组的线性相关性可以用线性组合或线性表示的语言来表述. 我们用下面的命题给出两种这样的表述.

命题 2.2 设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 为 m 维向量组，则下述三条等价：

(i) $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性相关；

(ii) 存在不全为零的数 $k_1,k_2,\cdots,k_s$，使得

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = 0;$$

(iii) $\alpha_1,\alpha_2,\cdots,\alpha_s$ 中的某个向量可被其余的向量线性表示(这里 $s\geqslant 2$).

证 我们采用轮转证法.

(i)$\Longrightarrow$(ii). 将 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关的定义中的线性方程组改写为向量方程,则此线性方程组的非零解就是(ii)中所要求的 $k_1,k_2,\cdots,k_s$.

(ii)$\Longrightarrow$(iii). 设(ii)中的某个 $k_i\neq 0$. 移项(即在(ii)中的等式两端同时加上除 $k_i\alpha_i$ 之外的各项的负向量),得

$$k_i\alpha_i = -k_1\alpha_1 - \cdots - k_{i-1}\alpha_{i-1} - k_{i+1}\alpha_{i+1} - \cdots - k_s\alpha_s.$$

两端乘以 $1/k_i$,得

$$\alpha_i = -\frac{k_1}{k_i}\alpha_1 - \cdots - \frac{k_{i-1}}{k_i}\alpha_{i-1} - \frac{k_{i+1}}{k_i}\alpha_{i+1} - \cdots - \frac{k_s}{k_i}\alpha_s.$$

即 α_i 可被其余向量 $\alpha_1,\cdots,\alpha_{i-1},\alpha_{i+1},\cdots,\alpha_s$ 线性表示.

(iii)$\Longrightarrow$(i). 设 $\alpha_i=h_1\alpha_1+\cdots+h_{i-1}\alpha_{i-1}-h_{i+1}\alpha_{i+1}-\cdots-h_s\alpha_s$,其中 h_i 为数($i=1,\cdots,i-1,i+1,\cdots,s$). 移项,得

$$h_1\alpha_1 + \cdots + h_{i-1}\alpha_{i-1} - \alpha_i + h_{i+1}\alpha_{i+1} + \cdots + h_s\alpha_s = 0.$$

将此向量方程写成齐次线性方程组的形式,则此线性方程组有非零解

$$x_1 = h_1,\cdots,x_{i-1} = h_{i-1},x_i = -1,x_{i+1} = h_{i+1},\cdots,x_s = h_s.$$

所以 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性相关. ▌

这个命题中的(ii),(iii)两条作为线性相关的两个等价的定义,比起原来的定义是变得抽象了,因为它们完全没有涉及向量的坐标. 正因为如此,它们有更普遍的适用性,在第四章中我们将会看到这一点.

由于命题与其逆否命题等价,我们有下面的推论:

推论 设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 为 m 维向量组,则下述三条等价:

(i) $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关;

(ii) 若 $k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=0$($k_1,k_2,\cdots,k_s$ 为数),则

$$k_1 = k_2 = \cdots = k_s = 0;$$

(iii) $\alpha_1, \alpha_2, \cdots, \alpha_s$ 中任一向量不能被其余的向量线性表出(这里 $s \geqslant 2$).

此推论中的(ii)和(iii)两条是线性无关的较为抽象的两个等价定义. 其中的(ii)在证明向量组线性无关时最为有用,但其逻辑结构比较复杂,比较难于掌握. 正确地理解和运用这个定义的关键是记住:(ii)的等价说法是"向量方程

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_s\alpha_s = 0$$

只有零解".

仅由一个向量构成的向量组($s=1$)是否线性相关取决于该向量是否为零向量. 这与命题 2.2 的(i),(ii)两条以及其推论中的(i),(ii)两条相符合,尽管命题及推论中的两个第(iii)条在这种情形下不再有意义.

例 4 给定 m 维向量组 α, 0, β,此向量组线性相关. 这是因为:取

$$k_1 = 0,\ k_2 = 1,\ k_3 = 0,$$

这是不全为零的一组数. 我们有

$$k_1\alpha + k_2 \cdot 0 + k_3\beta = 0 \cdot \alpha + 1 \cdot 0 + 0 \cdot \beta = 0.$$

从例 4 立刻可以看出,一个向量组中如包含有一个零向量,那么它必线性相关. 特别地,仅由一个零向量组成的向量组必线性相关. 不难看出,由一个非零向量 α 组成的向量组是线性无关的.

例 5 给定向量组 α, $-\alpha$, β(其中 α, β 为任意两个同维数的向量),此向量组线性相关. 这是因为:取

$$k_1 = 1,\ k_2 = 1,\ k_3 = 0,$$

这是一组不全为零的数,有

$$k_1\alpha + k_2(-\alpha) + k_3\beta = \alpha + (-\alpha) + 0 \cdot \beta = 0.$$

例 6 向量组 $k\alpha, \alpha, \beta$ 线性相关. 因为其中第 1 个向量能被其余两个向量线性表示

$$(k\alpha) = k \cdot \alpha + 0 \cdot \beta.$$

例 7 给定向量组

$$\alpha_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 2 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 2 \\ -1 \\ 3 \\ 1 \end{pmatrix},$$

判断它是否线性相关.

解 考察向量方程

$$x_1\alpha_1 + x_2\alpha_2 = \begin{pmatrix} x_1 \\ -x_1 \\ 0 \\ 2x_1 \end{pmatrix} + \begin{pmatrix} 2x_2 \\ -x_2 \\ 3x_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ -x_1 - x_2 \\ 3x_2 \\ 2x_1 + x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

两向量相等,则分量相同,故得

$$x_1 + 2x_2 = 0, \quad -x_1 - x_2 = 0, \quad 3x_2 = 0, \quad 2x_1 + x_2 = 0.$$

显然有 $x_1=0$, $x_2=0$. 即不存在不全为零的数 k_1, k_2, 使 $k_1\alpha_1+k_2\alpha_2=0$. 于是 α_1,α_2 线性无关.

最后再来举一个重要的例子.

例 8 给定如下 n 个 n 维向量

$$\varepsilon_1 = (1, 0, 0, \cdots, 0),$$
$$\varepsilon_2 = (0, 1, 0, \cdots, 0),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\varepsilon_n = (0, 0, \cdots, 0, 1),$$

称之为 n 维向量空间的 n 个**单位向量**. 我们来证明 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 线性无关. 若

$$k_1\varepsilon_1 + k_2\varepsilon_2 + \cdots + k_n\varepsilon_n = 0,$$

则因为

$$k_1\varepsilon_1 + k_2\varepsilon_2 + \cdots + k_n\varepsilon_n = (k_1, k_2, \cdots, k_n) = 0,$$

故必有 $k_1=k_2=\cdots=k_n=0$,因而向量组 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 线性无关.

向量组的秩

定义 给定两个向量组

$$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r, \tag{Ⅰ}$$

$$\beta_1,\ \beta_2,\ \cdots,\ \beta_s, \tag{Ⅱ}$$

如果向量组(Ⅱ)中每一个向量都能被向量组(Ⅰ)线性表示,反过来,向量组(Ⅰ)中每个向量也都能被向量组(Ⅱ)线性表示,则称向量组(Ⅰ)和向量组(Ⅱ)**线性等价**.

例 9 给定两个向量组

$$\alpha,\ \alpha,\ \beta, \tag{Ⅰ}$$

$$\alpha,\ \beta, \tag{Ⅱ}$$

则(Ⅰ)与(Ⅱ)线性等价. 这是因为:

(i) (Ⅱ)中每个向量能被(1)线性表示:

$\alpha = 1\cdot\alpha + 0\cdot\alpha + 0\cdot\beta$; $\beta = 0\cdot\alpha + 0\cdot\alpha + 1\cdot\beta$.

(ii) (Ⅰ)中每个向量也能被(Ⅱ)线性表示:

$$\alpha = 1\cdot\alpha + 0\cdot\beta;\quad \beta = 0\cdot\alpha + 1\cdot\beta.$$

在上述例子中,(Ⅱ)实际上是(Ⅰ)的一个部分组,是剔除(Ⅰ)中的“多余”向量所得到的. 这个简单的例子说明引入向量组线性等价概念的重要意义. 因为给定一个向量组,其中很可能包含有“多余”的向量. 于是,我们自然就想把这些“多余”的向量剔除掉,也就是设法用一个与它线性等价的部分组来取代它,而这个部分组中不再包含有“多余”的向量. 这就是本段落所要解决的问题.

为说明向量组线性等价概念的基本性质,我们先来证明一个命题.

命题 2.3 给定两个向量组

$$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r, \tag{Ⅰ}$$

$$\beta_1,\ \beta_2,\ \cdots,\ \beta_s, \tag{Ⅱ}$$

且(Ⅱ)中每一个向量 β_i 均能被向量组(Ⅰ)线性表示. 那么,当向量 γ 可被向量组(Ⅱ)线性表示时,它也就能被向量组(Ⅰ)线性表

示.

证 设 $\gamma=l_1\beta_1+l_2\beta_2+\cdots+l_s\beta_s$. 又设

$$\beta_1 = k_{11}\alpha_1 + k_{12}\alpha_2 + \cdots + k_{1r}\alpha_r,$$
$$\beta_2 = k_{21}\alpha_1 + k_{22}\alpha_2 + \cdots + k_{2r}\alpha_r,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\beta_s = k_{s1}\alpha_1 + k_{s2}\alpha_2 + \cdots + k_{sr}\alpha_r.$$

以 $l_1,l_2,\cdots,l_s$ 分别乘上面的第 1,2,…,s 个等式,然后相加起来,得

$$\gamma = \sum_{i=1}^{s} l_i\beta_i = \left(\sum_{i=1}^{s} l_ik_{i1}\right)\alpha_1 + \left(\sum_{i=1}^{s} l_ik_{i2}\right)\alpha_2 + \cdots + \left(\sum_{i=1}^{s} l_ik_{is}\right)\alpha_r.$$

即 γ 可被 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性表示. ▌

向量组线性等价的概念具有下列三条基本性质:

(i) **反身性** 每个向量组和它自身线性等价;

(ii) **对称性** 如果(Ⅰ)和(Ⅱ)线性等价,则(Ⅱ)与(Ⅰ)线性等价;

(iii) **传递性** 如果(Ⅰ)和(Ⅱ)线性等价,(Ⅱ)和(Ⅲ)线性等价,则(Ⅰ)和(Ⅲ)线性等价.

性质(i),(ii)是显然的. 性质(iii)可由命题 2.3 推得:因(Ⅲ)中每个向量能被(Ⅱ)线性表示,(Ⅱ)又被(Ⅰ)线性表示,故(Ⅲ)中每个向量能被(Ⅰ)线性表示. 反之,按同样理由,(Ⅰ)中每个向量也能被(Ⅲ)线性表示,故(Ⅰ)与(Ⅲ)线性等价.

下面给出本段落的主要概念.

定义 给定向量组

$$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_s, \qquad (Ⅰ)$$

如果它的一个部分组

$$\alpha_{i_1},\ \alpha_{i_2},\ \cdots,\ \alpha_{i_r} \qquad (Ⅱ)$$

满足如下两个条件:

(i) 向量组(Ⅰ)中每个向量都能被(Ⅱ)线性表示;

(ii) 向量组(Ⅱ)线性无关.

则称向量组(Ⅱ)是向量组(Ⅰ)的**极大线性无关部分组**.

显然,上面定义中的向量组(Ⅰ)和(Ⅱ)是线性等价的.因为:(Ⅱ)是(Ⅰ)的部分组,(Ⅱ)中每个向量当然都能被(Ⅰ)线性表示,而条件(i)又保证了(Ⅰ)中每个向量都能被(Ⅱ)线性表示.所以,一个向量组的极大线性无关部分组的实质是:从原向量组中挑出一部分来组成一个新向量组,使新向量组与原向量组线性等价,从而在一定意义下可用新向量组来代表原向量组.另一方面,新向量组线性无关,即其中没有"多余"的向量(因为这时其中任一向量都不能被其余向量线性表示,因而这些向量互相之间没有依赖关系,是互相"独立"的).

例 10 给定向量组

$$\alpha_1 = (1,0,0), \quad \alpha_2 = (0,1,0),$$
$$\alpha_3 = (0,0,1), \quad \alpha_4 = (1,1,1),$$
$$\alpha_5 = (1,1,0),$$

则 $\alpha_1,\alpha_2,\alpha_3$ 是它的一个极大线性无关部分组.因为:

(i) 所给向量组能由 $\alpha_1,\alpha_2,\alpha_3$ 线性表出

$$\alpha_1 = 1\cdot\alpha_1 + 0\cdot\alpha_2 + 0\cdot\alpha_3;$$
$$\alpha_2 = 0\cdot\alpha_1 + 1\cdot\alpha_2 + 0\cdot\alpha_3;$$
$$\alpha_3 = 0\cdot\alpha_1 + 0\cdot\alpha_2 + 1\cdot\alpha_3;$$
$$\alpha_4 = \alpha_1 + \alpha_2 + \alpha_3;$$
$$\alpha_5 = \alpha_1 + \alpha_2 + 0\cdot\alpha_3.$$

(ii) $\alpha_1,\alpha_2,\alpha_3$ 线性无关,这从本节例 8 即知.

请读者自己验证:向量组 $\alpha_1,\alpha_3,\alpha_4$ 也是它的一个极大线性无关部分组.由此可知,一个向量组的极大线性无关部分组不是唯一的.于是产生了一个问题:同一个向量组的不同的极大线性无关部分组中向量的个数是不是总是一样多呢?下面来回答这个问题.先证明几个命题.

命题 2.4 给定两个向量组

$$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r, \qquad (\text{Ⅰ})$$

$$\beta_1,\ \beta_2,\ \cdots,\ \beta_s, \tag{Ⅱ}$$

如果向量组(Ⅰ)中每个向量都能被(Ⅱ)线性表示，且 $r>s$，则向量组(Ⅰ)线性相关.

证 设

$$\begin{aligned}
\alpha_1 &= a_{11}\beta_1 + a_{12}\beta_2 + \cdots + a_{1s}\beta_s, \\
\alpha_2 &= a_{21}\beta_1 + a_{22}\beta_2 + \cdots + a_{2s}\beta_s, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\alpha_r &= a_{r1}\beta_1 + a_{r2}\beta_2 + \cdots + a_{rs}\beta_s.
\end{aligned}$$

分别用 $x_1,\cdots,x_2,x_r$ 乘上面第 1,2,…,r 个等式，然后相加，得

$$\sum_{m=1}^{r} x_m\alpha_m = \left(\sum_{m=1}^{r} a_{m1}x_m\right)\beta_1 + \left(\sum_{m=1}^{r} a_{m2}x_m\right)\beta_2 + \cdots + \left(\sum_{m=1}^{r} a_{ms}x_m\right)\beta_s.$$

考察齐次线性方程组

$$\begin{cases}
\sum\limits_{m=1}^{r} a_{m1}x_m = a_{11}x_1 + a_{21}x_2 + \cdots + a_{r1}x_r = 0, \\
\sum\limits_{m=1}^{r} a_{m2}x_m = a_{12}x_1 + a_{22}x_2 + \cdots + a_{r2}x_r = 0, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\sum\limits_{m=1}^{r} a_{ms}x_m = a_{1s}x_1 + a_{2s}x_2 + \cdots + a_{rs}x_r = 0,
\end{cases}$$

它有 s 个方程，r 个未知量，因为 $s<r$，由命题 1.2，它必有一组非零解

$$x_1 = k_1,\quad x_2 = k_2,\quad \cdots,\quad x_r = k_r.$$

此时有

$$\sum_{m=1}^{r} k_m\alpha_m = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_r\alpha_r = 0.$$

故 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性相关. ▌

命题 2.5 给定向量组

$$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_s, \tag{Ⅰ}$$

设它的某一个极大线性无关部分组为

$$\alpha_{i_1},\ \alpha_{i_2},\ \cdots,\ \alpha_{i_r}, \qquad (\mathrm{I}')$$

又有另一个向量组

$$\beta_1,\ \beta_2,\ \cdots,\ \beta_t, \qquad (\mathrm{II})$$

设它的某一个极大线性无关部分组为

$$\beta_{j_1},\ \beta_{j_2},\ \cdots,\ \beta_{j_l}. \qquad (\mathrm{II}')$$

如果(Ⅰ)与(Ⅱ)线性等价,则 $r=l$.

证 因为(Ⅰ′)与(Ⅰ)线性等价,(Ⅰ)与(Ⅱ)线性等价,(Ⅱ)与(Ⅱ′)线性等价,由线性等价关系的传递性知,(Ⅰ′)与(Ⅱ′)线性等价.(Ⅰ′)可被(Ⅱ′)线性表示且(Ⅰ′)线性无关,由命题 2.4 知 $r\leqslant l$. 反过来,(Ⅱ′)又被(Ⅰ′)线性表示,(Ⅱ′)也是线性无关的,故 $l\leqslant r$. 由此即得 $r=l$. ▌

推论 1 一个向量组的任意两个极大线性无关部分组中包含的向量个数相同.

证 在命题 2.5 中,取(Ⅰ)与(Ⅱ)为同一向量组即可. ▌

由推论 1,可给出如下重要概念.

定义 一个向量组的极大线性无关部分组中包含的向量个数称为该向量组的**秩**. 全由零向量组成的向量组的秩为零.

例 10 所给的向量组的秩为 3.

推论 2 两个线性等价的向量组的秩相等.

在本节的最后,我们再来指出一个重要的事实:如果检查一下上面几个命题的证明过程,不难看出,它们只以命题 2.1 为基础,而不依赖于向量的具体坐标表达式. 由于这一点,使得我们有可能对 n 维向量空间的概念从理论上进一步抽象化. 我们将在第四章中来做这个工作.

*集合内的等价关系

设 A 是一个非空的集合. 如果在 A 的元素之间定义了一种关系,记作"~",满足如下条件:

(i) 反身性，即对任意 $a\in A$，有 $a\sim a$；

(ii) 对称性，即若 $a\sim b$，则 $b\sim a$；

(iii) 传递性，即若 $a\sim b$，$b\sim c$，则 $a\sim c$.

则称此关系为 A 内的一个**等价关系**.

例如，在平面上全体三角形组成的集合中，"相似"是一个等价关系. 这是因为：任一三角形与自己相似(反身性)；若$\triangle ABC$ 相似于$\triangle A'B'C'$，则$\triangle A'B'C'$ 相似于$\triangle ABC$(对称性)；若$\triangle ABC$ 相似于$\triangle A'B'C'$，$\triangle A'B'C'$ 又相似于$\triangle A''B''C''$，则$\triangle ABC$ 相似于$\triangle A''B''C''$(传递性). 显然，在三角形集合中，"全等"也是一个等价关系. 又例如，在整数集合中，"两个数的差是 3 的倍数"具有上述三条性质：以 a,b,c 表示任意三个整数，则 $a-a=0=0\times 3$(反身性)；$a-b=m\times 3$(m 为整数)，则 $b-a=(-m)\times 3$(对称性)；$a-b=m\times 3$，$b-c=n\times 3$，则 $a-c=(m+n)\times 3$(传递性). 再例如：n 元线性方程组构成的集合中，"同解"是一个等价关系.

与等价关系密切相关的是等价类的概念. 确切地说，设 S 是一个集合，"$\sim$"是 S 中的任一等价关系，a 为 S 的任一元素，则把 S 中与 a 等价的所有元素构成的子集合称为 a 所在的**等价类**. 如果用 $\bar{a}$ 表示 a 所在的等价类，即有

$$\bar{a}=\{s\in S\,|\,s\sim a\}.$$

一个简单的事实是：

$$a\sim b \text{ 的充分必要条件是 } \bar{a}=\bar{b}\quad (\forall\ a,b\in S).$$

我们先说明必要性. 设 $a\sim b$，则对任一 $s\in\bar{a}$，由等价类的定义，有 $s\sim a$. 而 $a\sim b$，由传递性，有 $s\sim b$，此即 $s\in\bar{b}$，所以 $\bar{a}\subseteq\bar{b}$. 反之，由等价关系的对称性，有 $b\sim a$. 由同样的推理得知 $\bar{b}\subseteq\bar{a}$. 所以 $\bar{a}=\bar{b}$. 再说明充分性. 设 $\bar{a}=\bar{b}$，由等价类的定义立即得到 $a\in\bar{a}=\bar{b}$，进而 $a\sim b$. 这就完成了上述事实的证明.

等价关系的重要性在于：对于任一集合 S 中的任一给定的等价关系，S 等于所有等价类的并集，而且不同的等价类没有公共元素. 换句话说，S 等于所有等价类的无交并. 这是因为：S 显然包含

所有等价类的并集；而 S 中任一元素都属于它所在的等价类，所以 S 含于所有等价类的并集. 这就证明了 S 等于所有等价类的并集. 再有，如果等价类 $\bar{a}$ 与 $\bar{b}$ 有公共元素 c，则由上面刚刚证明的事实知，$\bar{a}=\bar{c}=\bar{b}$. 这说明：如果 $\bar{a}\neq\bar{b}$，则 $\bar{a}$ 与 $\bar{b}$ 必无公共元素.

数学中的一个重要方法是：为了研究某个集合的某一问题，在此集合中引入相应的等价关系，然后寻求各等价类中形式最简单、性质最好的元素，从而使问题的研究得以简化. 例如，本节刚刚引入的线性等价是 m 维向量组构成的集合中的一个等价关系(注意：此集合中的每个元素都是一个向量组，而不是单个的向量). 上面定义的"极大线性无关部分组"就是相应于这个等价关系的等价类中最简单的元素. 数学中的这个研究方法将在本书以下各章中反复地得到体现.

习 题 二

1. 将习题一第 1 题中各线性方程组改写成向量方程(要求具体写出所涉及的向量).

2. 将下列向量方程改写为普通方程，并求解：

(1) 设

$$\alpha_1=\begin{pmatrix}-1\\0\\2\\1\\1\end{pmatrix},\quad \alpha_2=\begin{pmatrix}0\\0\\1\\1\\2\end{pmatrix},\quad \alpha_3=\begin{pmatrix}7\\-2\\1\\-1\\0\end{pmatrix},\quad \beta=\begin{pmatrix}1\\2\\-3\\1\\0\end{pmatrix},$$

且

$$x_1\alpha_1+x_2\alpha_2+x_3\alpha_3=\beta.$$

(2) 设

$$\alpha_1=\begin{pmatrix}\sqrt{2}\\-\sqrt{3}\\\mathrm{i}\\1\end{pmatrix},\quad \alpha_2\begin{pmatrix}\mathrm{i}\\-\mathrm{i}\\\sqrt{3}\\0\end{pmatrix},\quad \alpha_3=\begin{pmatrix}0\\\mathrm{i}\\-\mathrm{i}\\\sqrt{2}\end{pmatrix},\quad \beta=\begin{pmatrix}1\\\mathrm{i}\\0\\0\end{pmatrix},$$

且 $$x_1\alpha_1+x_2\alpha_2+x_3\alpha_3=\beta.$$

3. 给定向量

$\alpha_1=(0,-1,1,1)$， $\alpha_2=(1,-1,1,2)$，

$\alpha_3=(0,1,1,2)$， $\alpha_4=(2,2,1,3)$，

$\alpha_5=(0,-1,-1,-1)$，$\beta=(3,1,4,8)$，

将 β 表成 $\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5$ 的线性组合.

4. 将向量 β 表成向量 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 的线性组合.

(1) $\alpha_1=(1,1,1,1)$， $\alpha_2=(1,1,-1,-1)$，

$\alpha_3=(1,-1,1,-1)$， $\alpha_4=(1,-1,-1,1)$，

$\beta=(1,2,1,1)$.

(2) $\alpha_1=(1,1,0,1)$， $\alpha_2=(2,1,3,1)$，

$\alpha_3=(1,1,0,0)$， $\alpha_4=(0,1,-1,-1)$，

$\beta=(0,0,0,1)$.

5. 判断下列向量组是否线性相关.

(1) $\alpha_1=(1,3,5,-4,0)$， $\alpha_2=(1,3,2,-2,1)$，

$\alpha_3=(1,-2,1,-1,-1)$，$\alpha_4=(1,-4,1,1,-1)$.

(2) $\alpha_1=(1,-2,3,-4)$， $\alpha_2=(0,1,-1,1)$，

$\alpha_3=(1,3,0,1)$， $\alpha_4=(0,-7,3,1)$.

(3) $\alpha_1=(1,2,3,-1)$， $\alpha_2=(3,2,1,-1)$，

$\alpha_3=(2,3,1,1)$， $\alpha_4=(2,2,2,-1)$，

$\alpha_5=(5,5,2,0)$.

(4) $\alpha_1=(1,-1,0,0,0)$， $\alpha_2=(0,1,-1,0,0)$，

$\alpha_3=(0,0,1,-1,0)$， $\alpha_4=(0,0,0,1,-1)$，

$\alpha_5=(-1,0,0,0,1)$.

6. 证明向量组 $\alpha_1,\alpha_2,\alpha_3$ 与向量组 $\alpha_1+\alpha_2,\alpha_2+\alpha_3,\alpha_3+\alpha_1$ 线性等价.

7. 设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关，证明 $\alpha_1,\alpha_1+\alpha_2,\cdots,\alpha_1+\alpha_2+\cdots+\alpha_s$ 线性无关.

8. 证明：如果向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关，而 $\alpha_1,\alpha_2,\cdots,\alpha_s,\beta$

线性相关，则 β 可被向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表示.

9. 证明：一个线性无关向量组的任一个部分组也线性无关.

10. 证明：如果一个向量组有一个部分组线性相关，那么该向量组也线性相关.

11. 给定 n 维向量组

$$
\begin{aligned}
\alpha_1 &= (a_{11}, a_{12}, \cdots, a_{1n}), \\
\alpha_2 &= (a_{21}, a_{22}, \cdots, a_{2n}), \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\alpha_m &= (a_{m1}, a_{m2}, \cdots, a_{mn}),
\end{aligned}
$$

从每个向量中去掉第 $i_1,i_1,\cdots,i_s$ 个分量，得到一个 $n-s$ 维的新向量组 $\alpha_1',\alpha_2',\cdots,\alpha_m'$. 证明：

（1） 若 $\alpha_1',\alpha_2',\cdots,\alpha_m'$ 线性无关，则 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 也线性无关；

（2） $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性相关，则 $\alpha_1',\alpha_2',\cdots,\alpha_m'$ 也线性相关.

12. 给定 n 维向量组

$$
\begin{aligned}
\alpha_1 &= (a_{11}, a_{12}, \cdots, a_{1n}), \\
\alpha_2 &= (a_{21}, a_{22}, \cdots, a_{2n}), \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\alpha_m &= (a_{m1}, a_{m2}, \cdots, a_{mn}), \\
\beta &= (b_1, b_2, \cdots, b_n),
\end{aligned}
$$

对它们的分量做如下变化

（1） 把第 i 个分量与第 j 个分量互换；

（2） 把第 i 个分量乘以非零常数 c；

（3） 把第 j 个分量加上第 i 个分量的 k 倍.

设经过上述三种变换的任一种后得到新向量组 $\alpha_1',\alpha_2',\cdots,\alpha_m',\beta'$，证明：

（1） 若 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性无关，则 $\alpha_1',\alpha_2',\cdots,\alpha_m'$ 也线性无关；

（2） 若 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性相关，则 $\alpha_1',\alpha_2',\cdots,\alpha_m'$ 也线性相关；

（3） 若 β 能被 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性表示，则 β' 也能被 $\alpha_1',\alpha_2',\cdots,$

α_m' 线性表示.

13. 证明:向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s(s\geqslant2,\alpha_1\neq0)$线性相关的充分必要条件是,至少有一个 α_i 可被 $\alpha_1,\alpha_2,\cdots,\alpha_{i-1}$线性表示.

14. 求下列向量组的极大线性无关部分组和秩.

(1) $\alpha_1=(-1,0,1,0,0)$, $\alpha_2=(1,1,1,1,0)$,

$\alpha_3=(0,1,2,1,0)$.

(2) $\alpha_1=(1,0,0)$, $\alpha_2=(1,1,0)$,

$\alpha_3=(1,1,1)$, $\alpha_4=(0,1,1)$,

$\alpha_5=(0,0,1)$.

(3) $\alpha_1=(2,3,1,1)$, $\alpha_2=(4,6,2,2)$,

$\alpha_3=(0,1,2,1)$, $\alpha_4=(0,-1,-2,-1)$.

15. 证明:$\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$是向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的极大线性无关部分组当且仅当下述两条成立:

(i) $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$线性无关;

(ii) $\alpha_i,\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$线性相关,其中 α_i 为 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 中任一向量.

16. 已知 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的秩为 r,证明其中任意 r 个线性无关的向量都构成它的一个极大线性无关部分组.

17. 设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的秩为 r,而 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$是其中 r 个向量,使每个 $\alpha_i(i=1,2,\cdots,s)$都能被它们线性表示,证明 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$是 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的一个极大线性无关部分组.

18. 证明:如果向量组(Ⅰ)可以由向量组(Ⅱ)线性表示,那么(Ⅰ)的秩≤(Ⅱ)的秩.

19. 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是一个 n 维向量组,如果 n 维单位向量 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 可被它们线性表示,证明:$\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关.

20. 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是一个 n 维向量组,证明:$\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关的充分必要条件是,任一个 n 维向量都可被它们线性表示.

21. 证明一个向量组的任一线性无关部分组都可扩充成它

的一个极大线性无关部分组.

22. 设 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 与 $\alpha_1,\alpha_2,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_s$ 的秩相同,证明它们线性等价.

23. 设 $\beta_1=\alpha_2+\alpha_3+\cdots+\alpha_r,\beta_2=\alpha_1+\alpha_3+\cdots+\alpha_r,\cdots,\beta_r=\alpha_1+\alpha_2+\cdots+\alpha_{r-1}$. 证明:$\beta_1,\beta_2,\cdots,\beta_r$ 与 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 的秩相同.

§3 矩 阵 的 秩

在上一节中,我们在线性方程组和 n 维向量之间建立了联系,并讨论了 n 维向量空间的一些基本概念和性质. 下面我们把它们再应用到线性方程组理论中去. 我们知道,一个线性方程组可以用一个矩阵来代表,因此,我们把矩阵当作一个中间桥梁. 首先把上一节的结果应用于矩阵,然后再把在矩阵中所获得的结果应用于线性方程组.

给定一个 $m\times n$ 矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \tag{1}$$

它的每一列可以看作是一个 m 维向量,它有 n 个列,组成一个 m 维向量组,我们称之为矩阵 A 的**列向量组**. 同样,它的每一行可以看作一个 n 维向量,它有 m 个行,组成一个 n 维向量组,我们称之为矩阵 A 的**行向量组**.

定义 一个矩阵 A 的行向量组的秩称为 A 的**行秩**,它的列向量组的秩称为 A 的**的列秩**.

为着研究矩阵的行秩和列秩,我们在矩阵上定义初等变换如下:

定义 对矩阵 A 的行(列)施行如下变换:

(i) 互换两行(列)的位置;

(ii) 把某一行(列)乘以一个非零常数 c;

(iii) 把某一行(列)加上另一行(列)的 k 倍.

上述三种变换中的每一种都称为矩阵 A 的**初等行(列)变换**.

如果把矩阵 A 看作某个线性方程组的增广矩阵,那么它的行变换与方程组的初等变换一致.但解方程组时不允许做列初等变换,而矩阵本身可以做列初等变换.

注意矩阵的行(列)初等变换都是可逆的.因为:

(1) 如 A 经过互换 i,j 两行(列)变成矩阵 B,则 B 经互换 i,j 两行(列)就变回矩阵 A;

(2) 如 A 第 i 行(列)乘以非零常数 c 后变为矩阵 B,则 B 的第 i 行(列)乘以 $1/c$ 即变回矩阵 A;

(3) 如 A 的第 j 行(列)加上第 i 行(列)的 k 倍后变成矩阵 B,则 B 的第 j 行(列)加上第 i 行(列)的 $-k$ 倍后即变回矩阵 A.

由此可知:如果矩阵 A 经过若干次初等行、列变换化为矩阵 B,则 B 也可经过若干次初等行、列变换化为矩阵 A.

命题 3.1 矩阵 A 的行秩在行初等变换下保持不变;矩阵 A 的列秩在列初等变换下也保持不变.

证 只证行秩在行初等变换下不变,列秩的证法相同,不必重复.

设 A 的行向量组为 $\alpha_1,\alpha_2,\cdots,\alpha_m$.

(i) 互换 A 的 i,j 两行相当于把 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 中 α_i 与 α_j 两向量调换一下位置,所得新向量组显然与原向量组线性等价;

(ii) 把 A 的第 i 行乘以 $c\neq 0$,所得新矩阵的行向量组为 $\alpha_1,\cdots,c\alpha_i,\cdots,\alpha_m$.因为

$$\alpha_i = \frac{1}{c}\cdot(c\alpha_i),\quad (c\alpha_i) = c\cdot\alpha_i,$$

故新向量组与原向量组能互相线性表示,即线性等价;

(iii) 把 A 的第 j 行加上第 i 行的 k 倍后,所得新矩阵的行向

量组为

$$\alpha_1,\cdots,\alpha_i,\cdots,\alpha_j+k\alpha_i,\cdots,\alpha_m$$

(凡未写出的都与原来相同). 因为

$$\alpha_j=(\alpha_j+k\alpha_i)+(-k)\cdot\alpha_i,$$

$$(\alpha_j+k\alpha_i)=\alpha_j+k\cdot\alpha_i,$$

故新向量组与原向量组能互相线性表示,即线性等价.

根据命题 2.5 的推论 2,线性等价的向量组的秩相同,故矩阵 A 经过一次初等行变换后其行秩不变. 那么,经过任何次初等行变换后行秩也不会变化. ▌

在进一步讨论之前,先介绍一个概念. 把矩阵 A 的行与列互换之后,得到一个 $n\times m$ 矩阵

$$A'=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix},$$

A' 称为矩阵 A 的**转置矩阵**. 在本书中,我们固定用 A' 表示矩阵 A 的转置矩阵,下面不再重复说明. 注意 A' 中元素的下角标不再代表它所在的位置.

例如,设

$$A=\begin{pmatrix} -1 & 0 & 1 & 3 & -1 \\ 2 & 4 & 1 & 0 & 1 \\ 0 & 1 & -1 & 0 & 2 \end{pmatrix},$$

则其转置矩阵为

$$A'=\begin{pmatrix} -1 & 2 & 0 \\ 0 & 4 & 1 \\ 1 & 1 & -1 \\ 3 & 0 & 0 \\ -1 & 1 & 2 \end{pmatrix}.$$

因为矩阵 A 的行向量组变成 A' 的列向量组,所以 A 的行秩

等于 A' 的列秩. 同样, A 的列向量组变成 A' 的行向量组, 所以 A 的列秩等于 A' 的行秩.

命题 3.2 矩阵 A 的行秩在列初等变换下保持不变; 矩阵 A 的列秩在行初等变换下也保持不变.

证 分两步证明.

(1) 先证 A 的列秩在行初等变换下保持不变. 设 A 的列向量组为 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 其列秩为 r. 不妨设 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 为列向量组的一个极大线性无关部分组. 假定 A 经初等行变换后所得新矩阵的列向量组为 $\alpha_1', \alpha_2', \cdots, \alpha_n'$, 我们只要证 $\alpha_1', \alpha_2', \cdots, \alpha_r'$ 是它的一个极大线性无关部分组就可以了.

(i) 先证 $\alpha_1', \alpha_2', \cdots, \alpha_r'$ 线性无关. 以 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 为列向量排成一矩阵 B. 因 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 线性无关, 故以 B 为系数矩阵的齐次线性方程组只有零解; 另一方面, 以 $\alpha_1', \alpha_2', \cdots, \alpha_r'$ 为列向量排成矩阵 B_1, B_1 可由 B 经初等行变换得到. 由命题 1.1, 以 B_1 为系数矩阵的齐次线性方程组和以 B 为系数矩阵的齐次线性方程组同解, 因而也只有零解. 故 $\alpha_1', \alpha_2', \cdots, \alpha_r'$ 线性无关.

(ii) 再证任一 α_i' 可被 $\alpha_1', \alpha_2', \cdots, \alpha_r'$ 线性表示. 为此, 考察以 $\alpha_1, \alpha_2, \cdots, \alpha_r, \alpha_i$ 为列向量组的矩阵 $\overline{B}$. 因 α_i 可被 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 线性表示, 故以 $\overline{B}$ 为增广矩阵的线性方程组有解; 另一方面, 以 $\alpha_1', \alpha_2', \cdots, \alpha_r', \alpha_i'$ 为列向量组的矩阵 $\overline{B}_1$ 可由 $\overline{B}$ 经初等行变换得到. 由命题 1.1, 以 $\overline{B}_1$ 为增广矩阵的线性方程组和以 $\overline{B}$ 为增广矩阵的线性方程组同解, 因而也是有解的. 这说明 α_i' 可被 $\alpha_1', \alpha_2', \cdots, \alpha_r'$ 线性表示.

(2) 现在证 A 的行秩在列初等变换下不变. 考察 A'. A 的行向量组是 A' 的列向量组, 对 A 做列初等变换等价于对 A' 做行初等变换. 根据(1), A' 的列秩在行初等变换下不变, 这说明 A 的行秩在列初等变换下不变. ▌

综合命题 3.1 和 3.2, 矩阵 A 的行秩在行和列的初等变换下

都保持不变，其列秩也有同样的性质. 我们来看一看，同时对 A 做行和列的初等变换能把它变成什么样子.

我们假定 A 的元素不全为零(否则，A 的行秩和列秩都是零，这可以先排除在外).

(1) 在矩阵 A 中，如果 $a_{11}=0$，我们就在矩阵中找一个不为零的元素，设为 a_{ij}. 先对换 $1,i$ 两行，再对换 $1,j$ 两列，这就把 a_{ij} 调换到第 1 行第 1 列的位置上. 所以我们总可以假定 $a_{11}\neq 0$.

(2) 若 $a_{11}\neq 0$，利用初等行变换把 A 变成如下形状

$$A \longrightarrow \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & b_{m2} & & b_{mn} \end{bmatrix},$$

然后再利用初等列变换把 A 进一步变为

$$A \longrightarrow \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & b_{22} & b_{23} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & b_{m2} & b_{m3} & \cdots & b_{mn} \end{bmatrix}.$$

如此继续对右下角的 $(m-1)\times(n-1)$ 矩阵重复上述步骤.

经过连续施行上述初等行、列变换之后，矩阵 A 可变成下列三种阶梯形之一

$$\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & 0 & & \\ & & 1 & & & & \\ & & & 0 & & & \\ & 0 & & & \ddots & & \\ & & & & & 0 \cdots & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & & & & \\ & \ddots & & 0 & \\ & & 1 & & \\ & & & 0 & \\ & 0 & & & \ddots \\ & & & & & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & & & \\ & \ddots & & 0 \\ & & 1 & \\ & & & 0 \\ & & & & \ddots \\ & & & & & 0 \\ & 0 & & & & \vdots \\ & & & & & 0 \end{bmatrix}.$$

($n>m$ 时)　　　　($n=m$ 时)　　　　($n<m$ 时)

上述三种阶梯形矩阵称为 A 的**标准形**. 设标准形中 1 的个数为 r，

则标准形的行秩和列秩都是 r. 这是因为:它的前 r 个行向量即为其行向量组的极大线性无关部分组;它的前 r 个列向量也是其列向量组的极大线性无关部分组.

由命题 3.1 和 3.2,A 的行秩等于其标准形的行秩,A 的列秩等于其标准形的列秩,而标准形的行秩和列秩相等. 由此,我们得到如下一个重要结论:

命题 3.3 矩阵的行秩等于列秩.

定义 一个矩阵 A 的行秩或列秩称为该矩阵的**秩**,记作 $r(A)$.

因此,$r(A)$既是 A 的行向量组的极大线性无关部分组中向量的个数,也是 A 的列向量组的极大线性无关部分组中向量的个数. 从命题 3.1 和 3.2 可以得到矩阵秩的如下计算方法:利用初等行变换或列变换把它化为阶梯形(不必化为标准形),容易看出,阶梯形矩阵的秩就等于其阶梯数. 如果要求向量组的秩,可以把它按横的方式排成一个矩阵,也可按竖的方式排成一个矩阵,然后计算矩阵的秩就可以了.

例 1 求下面向量组的秩

$$\alpha_1=(1,-1,0,1,1),\quad \alpha_2=(2,-2,0,2,2),$$
$$\alpha_3=(1,1,1,0,0),\quad \alpha_4=(2,0,1,1,1).$$

解 把它们作为行排成 4×5 矩阵,再用初等变换将矩阵化为阶梯形

$$\begin{pmatrix}1&-1&0&1&1\\2&-2&0&2&2\\1&1&1&0&0\\2&0&1&1&1\end{pmatrix}\longrightarrow\begin{pmatrix}1&-1&0&1&1\\0&0&0&0&0\\0&2&1&-1&-1\\0&2&1&-1&-1\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1&-1&0&1&1\\0&2&1&-1&-1\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}.$$

最后阶梯形矩阵的秩为 2，故原矩阵秩为 2，因而向量组的秩也是 2. 如果再使用列初等变换，最后的阶梯矩阵又可以进一步化为如下标准形

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

它包含两个 1，秩为 2.

例 2 求下面向量组的极大线性无关部分组

$\alpha_1=(2,0,1,1)$，$\alpha_2=(-1,-1,-1,-1)$，

$\alpha_3=(1,-1,0,0)$，$\alpha_4=(0,-2,-1,-1)$.

解 在本例中要求的是极大线性无关部分组，而不只是秩. 我们采用如下办法：把这向量组作为行排成一个矩阵，同时把该向量的希腊字母写在它的右方

$$A=\begin{pmatrix} 2 & 0 & 1 & 1 & \alpha_1 \\ -1 & -1 & -1 & -1 & \alpha_2 \\ 1 & -1 & 0 & 0 & \alpha_3 \\ 0 & -2 & -1 & -1 & \alpha_4 \end{pmatrix}.$$

然后对上面的矩阵做行初等变换（现在不能做列初等变换了），在这过程中右边的用希腊字母标出的向量也跟着变. 这样，在变换过程中，每个行向量永远等于右边用希腊字母表示的向量.

$$A\to\begin{pmatrix} 1 & -1 & 0 & 0 & \alpha_3 \\ 2 & 0 & 1 & 1 & \alpha_1 \\ -1 & -1 & -1 & -1 & \alpha_2 \\ 0 & -2 & -1 & -1 & \alpha_4 \end{pmatrix}\to\begin{pmatrix} 1 & -1 & 0 & 0 & \alpha_3 \\ 0 & 2 & 1 & 1 & \alpha_1-2\alpha_3 \\ 0 & -2 & -1 & -1 & \alpha_2+\alpha_3 \\ 0 & -2 & -1 & -1 & \alpha_4 \end{pmatrix}$$

$$\longrightarrow\begin{pmatrix} 1 & -1 & 0 & 0 & \alpha_3 \\ 0 & 2 & 1 & 1 & \alpha_1-2\alpha_3 \\ 0 & 0 & 0 & 0 & \alpha_1+\alpha_2-\alpha_3 \\ 0 & 0 & 0 & 0 & \alpha_1-2\alpha_3+\alpha_4 \end{pmatrix}.$$

矩阵化成阶梯形后，可看出其秩为 2，故原向量组的秩为 2，其极大线性无关部分组向量个数为 2. 另一个方面，最后两个行向量为零，这表示

$$\alpha_1+\alpha_2-\alpha_3=0,\quad \alpha_1-2\alpha_3+\alpha_4=0.$$

从这两个向量方程解出两个向量

$$\alpha_3=\alpha_1+\alpha_2;\quad \alpha_4=\alpha_1+2\alpha_2.$$

于是整个向量组可被 α_1,α_2 线性表示，因而原向量组与 α_1,α_2 线性等价. 已知其秩为 2，那么 α_1,α_2 的秩也是 2，因而它线性无关. 这说明 α_1,α_2 即为原向量组的一个极大线性无关部分组.

上面所介绍的，是求一个向量组的极大线性无关部分组的一般性方法.

习 题 三

1. 计算下列矩阵的秩.

(1) $\begin{pmatrix}0&1&1&-1&2\\0&2&-2&-2&0\\0&-1&-1&1&1\\1&1&0&1&-1\end{pmatrix}$； (2) $\begin{pmatrix}1&-1&2&1&0\\2&-2&4&-2&0\\3&0&6&-1&1\\0&3&0&0&1\end{pmatrix}$；

(3) $\begin{pmatrix}14&12&6&8&2\\6&104&21&9&17\\7&6&3&4&1\\35&30&15&20&5\end{pmatrix}$； (4) $\begin{pmatrix}1&0&0&1&4\\0&1&0&2&5\\0&0&1&3&6\\1&2&3&14&32\\4&5&6&32&77\end{pmatrix}$；

(5) $\begin{pmatrix}1&0&1&0&0\\1&1&0&0&0\\0&1&1&0&0\\0&0&1&1&0\\0&1&0&1&1\end{pmatrix}$.

2. 求下面向量组的极大线性无关部分组和秩.

(1) $\alpha_1=(6,4,1,-1,2)$, $\alpha_2=(1,0,2,3,-4)$,

$\alpha_3=(1,4,-9,-16,22)$, $\alpha_4=(7,1,0,-1,3)$.

(2) $\alpha_1=(1,-1,2,4)$, $\alpha_2=(0,3,1,2)$,

$\alpha_3=(3,0,7,14)$, $\alpha_4=(1,-1,2,0)$,

$\alpha_5=(2,1,5,6)$.

3. 设 A,B 为 $m\times n$ 矩阵.

(1) 若以 $A\sim B$ 表示 A 可以经过行、列初等变换化为 B,证明 $\sim$ 是 $m\times n$ 矩阵的集合中的等价关系.

(2) 若以 $A\sim B$ 表示 A 可以经过行初等变换化为 B,证明 $\sim$ 是等价关系,并且 $r(A)=m$ 时 A 可单用行初等变换化为如下形式

$$\begin{pmatrix} 1 & & & 0 & \\ & 1 & & & \\ & & \ddots & & C \\ & 0 & & \ddots & \\ & & & & 1 \end{pmatrix},$$

其中 C 为 $m\times(n-m)$ 矩阵.

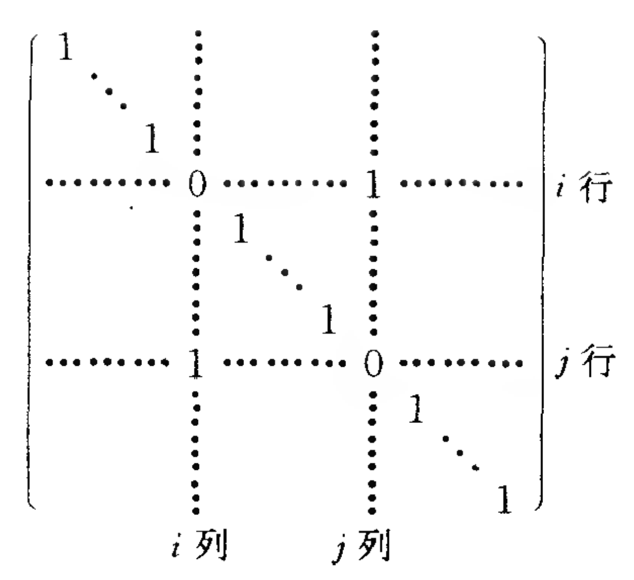
(3) 若以 $A\sim B$ 表示 A 可以经过列初等变换化为 B,证明 $\sim$ 也是等价关系,并且 $r(A)=n$ 时 A 可单用列初等变换化为如下形式

$$\begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & 0 & \ddots & \\ & & & 1 \\ & & C & \end{pmatrix}.$$

其中 C 为 $(m-n)\times n$ 矩阵.

4. 如果单用初等行变换把矩阵 A 化为阶梯形矩阵 F,证明 $r(A)$ 等于 F 中不为零的行向量的数目(即 F 的阶梯数).

5. 求下面 $n\times n$ 矩阵的秩



(空白处的元素均为零).

6. 证明:对任一矩阵 A,有 $r(A)=r(A')$.

7. 给定 $m\times n$ 矩阵 A,$m\times s$ 矩阵 B,把它们并排放在一起得 $m\times(n+s)$ 矩阵 $C=(AB)$. 证明

$$\max(r(A),r(B))\leqslant r(C)\leqslant r(A)+r(B).$$

8. 给定两个 $m\times n$ 矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},\quad B=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix},$$

令

$$C=\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{pmatrix}.$$

证明:$r(C)\leqslant r(A)+r(B)$.

9. 设矩阵 A 经初等行变换变为矩阵 B,以 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 和 $\beta_1,\beta_2,\cdots,\beta_n$ 分别代表 A,B 的列向量组. 证明:若对某一组数 k_1,$k_2,\cdots,k_n$ 有

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n=0,$$

那么

$$k_1\beta_1 + k_2\beta_2 + \cdots + k_n\beta_n = 0.$$

10. 设 A 是一个 $m\times n$ 矩阵，从中任取 s 行，得一 $s\times n$ 矩阵 B. 证明

$$r(B) \geqslant r(A) + s - m.$$

11. 设 A 是一个 $m\times n$ 矩阵，且 $r(A)=0$ 或 1. 证明存在数 $a_1,a_2,\cdots,a_m$； $b_1,b_2,\cdots,b_n$，使

$$A = \begin{pmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & & \vdots \\ a_mb_1 & a_mb_2 & \cdots & a_mb_n \end{pmatrix}.$$

§4 齐次线性方程组

在做了上面一系列准备工作之后，我们就可以着手来解决本章开头所提出的线性方程组理论的三个基本问题了. 这一节先来讨论齐次线性方程组. 它比较简单，但又是整个问题的核心部分，解决了它，其它部分即可迎刃而解.

齐次线性方程组的基础解系

考察齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0. \end{cases} \quad (1)$$

它的任意一组解

$$x_1 = k_1,\ x_2 = k_2,\ \cdots,\ x_n = k_n$$

可以看成一个 n 维向量 $(k_1,k_2,\cdots,k_n)$，称之为线性方程组(1)的一个**解向量**. 显然，非齐次线性方程组的一组解也同样可以看作一个

n 维向量. 因此，以后我们一般把线性方程组的一组解当作一个向量来处理.

齐次线性方程组(1)的解具有如下性质：

(i) 如果 $\eta_1=(k_1,k_2,\cdots,k_n)$，$\eta_2=(l_1,l_2,\cdots,l_n)$是方程组(1)的两个解向量，则

$$\eta_1+\eta_2=(k_1+l_1,k_2+l_2,\cdots,k_n+l_n)$$

也是方程组(1)的解向量；

(ii) 如果 $\eta=(k_1,k_2,\cdots,k_n)$是方程组(1)的一个解向量，则对任意数 k，有

$$k\eta=(kk_1,kk_2,\cdots,kk_n)$$

也是方程组(1)的解向量.

这两条性质只要直接代入方程组进行验证就可以了. 例如对性质(1)，有

$$\sum_{j=1}^{n}a_{ij}k_j=0\quad(i=1,2,\cdots,m);$$

$$\sum_{j=1}^{n}a_{ij}l_j=0\quad(i=1,2,\cdots,m).$$

两式相加，得

$$\sum_{j=1}^{n}a_{ij}k_j+\sum_{j=1}^{n}a_{ij}l_j=\sum_{j=1}^{n}a_{ij}(k_j+l_j)=0.$$

这表示 $\eta_1+\eta_2$ 是方程组(1)的一组解. 性质(2)请读者自行验证.

根据性质(i)，(ii)可知，如果 $\eta_1,\eta_2,\cdots,\eta_s$ 是方程组(1)的一组解向量，则它们的任意线性组合 $k_1\eta_1+k_2\eta_2+\cdots+k_s\eta_s$ 仍然是方程组(1)的一个解向量.

定义 齐次线性方程组(1)的一组解向量 $\eta_1,\eta_2,\cdots,\eta_s$ 如果满足如下条件

(i) $\eta_1,\eta_2,\cdots,\eta_s$ 线性无关；

(ii) 方程组(1)的任一解向量都可被 $\eta_1,\eta_2,\cdots,\eta_s$ 线性表示.

那么，就称 $\eta_1,\eta_2,\cdots,\eta_s$ 是齐次线性方程组(1)的一个**基础解系**.

如果把方程组(1)的全部解放在一起看成一个向量组(一般说,它包含无穷多个向量),那么,基础解系(如果存在的话)实际上就是它的一个极大线性无关部分组.从这个意义上说,基础解系可以作为方程组(1)的解的一个代表系.知道了基础解系,做它的全部线性组合,就得到方程组的全部解.这样,齐次线性方程组解的研究可以归结为对它的基础解系的研究.

应当指出:如果齐次线性方程组(1)只有零解,那么它就没有基础解系.但为着叙述上的方便,我们说这样的齐次线性方程组的基础解系包含零个向量.

为了研究方程组(1)的基础解系,先需要一个简单的命题.

命题 4.1 如果向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关,而向量 β 可被它线性表示,则表示法是唯一的.

证 设 β 有两种表示法

$$\begin{aligned}\beta &= k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s \\ &= l_1\alpha_1 + l_2\alpha_2 + \cdots + l_s\alpha_s,\end{aligned}$$

则有

$$(k_1 - l_1)\alpha_1 + (k_2 - l_2)\alpha_2 + \cdots + (k_s - l_s)\alpha_s = 0.$$

因为 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关,故必有

$$k_1 - l_1 = k_2 - l_2 = \cdots = k_s - l_s = 0.$$

即两个表示法相同. ▌

齐次线性方程组(1)的系数矩阵为

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

定理 1 齐次线性方程组(1)的基础解系存在,且任一基础解系中解向量个数为 $n-r$,其中 n 为未知量个数,而 r 为系数矩阵 A 的秩 $r(A)$.

证 因为方程组(1)的任意两个基础解系(如果有的话)是互

相等价的，因而秩相等. 它们又是线性无关的，秩即等于其向量个数，故任意两个基础解系中包含相同数目的向量. 因此，我们只要找出一个基础解系，其中包含 $n-r$ 个向量，定理就得证了.

设矩阵 A 的列向量组为 $\alpha_1,\alpha_2,\cdots,\alpha_n$. 如果 $r(A)=r=n$，即 A 的列向量组线性无关，则方程组(1)只有零解，其基础解系包含

$$n-r=n-n=0$$

个向量，定理成立.

下面设 $r<n$. 设 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 为 A 的列向量组的一个极大线性无关部分组. 把方程组(1)写成向量形式

$$x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n=0. \tag{2}$$

因为 $\alpha_{r+1},\alpha_{r+2},\cdots,\alpha_n$ 均能被 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性表示，按命题 2.3，它们的任一线性组合也能被 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性表示. 再由命题 4.1，表示法是唯一的. 因此，任给 $x_{r+1},x_{r+2},\cdots,x_n$ 一组数值

$$x_{r+1}=k_{r+1},\ x_{r+2}=k_{r+2},\ \cdots,\ x_n=k_n,$$

则因 $\beta=-(k_{r+1}\alpha_{r+1}+\cdots+k_n\alpha_n)$ 能唯一地表成 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 的线性组合，所以存在唯一的一组数 $k_1,\cdots,k_r$ 使

$$k_1\alpha_1+\cdots+k_r\alpha_r+k_{r+1}\alpha_{r+1}+\cdots+k_n\alpha_n=0.$$

这说明：

(i) 方程(2)中未知量 $x_{r+1},\cdots,x_n$ 任取一组数值，都可唯一确定未知量 $x_1,\cdots,x_r$ 的一组值，从而得到方程组(1)的一组解；

(ii) 方程组的两组解 η_1,η_2，如它们在 $x_{r+1},\cdots,x_n$ 处取的值相同

$$\eta_1=(*,\cdots,*,k_{r+1},\cdots,k_n);$$
$$\eta_2=(*,\cdots,*,k_{r+1},\cdots,k_n)$$

(其中星号 * 表示该处数值无需具体标出)，则 $\eta_1=\eta_2$.

未知量 $x_{r+1},\cdots,x_n$ 称为方程组的**自由未知量**. 如果让这 $n-r$ 个自由未知量中某一个取值 1，其余取值零，就得到方程组(1)的一组解. 这样一共可得 $n-r$ 个解

$$\begin{aligned}\eta_1 &= (*,\cdots,*,1,0,\cdots,0),\\ \eta_2 &= (*,\cdots,*,0,1,\cdots,0),\\ &\cdots\cdots\cdots\cdots\cdots\cdots\\ \eta_{n-r} &= (*,\cdots,*,0,\cdots,0,1).\end{aligned} \tag{3}$$

我们来证明 $\eta_1,\eta_2,\cdots,\eta_{n-r}$是方程组(1)的一个基础解系.

(i) 先证 $\eta_1,\eta_2,\cdots,\eta_{n-r}$线性无关. 因(利用(3)式)

$$k_1\eta_1 + k_2\eta_2 + \cdots + k_{n-r}\eta_{n-r} = (*,\cdots,*,k_1,k_2,\cdots,k_{n-r}) = 0,$$

故
$$k_1=k_2=\cdots=k_{n-r}=0.$$

按定义知 $\eta_1,\eta_2,\cdots,\eta_{n-r}$线性无关.

(ii) 设 $\eta=(k_1,\cdots,k_r,k_{r+1}\cdots,k_n)$为方程组(1)的任一组解,命

$$\begin{aligned}\eta' &= k_{r+1}\eta_1 + k_{r+2}\eta_2 + \cdots + k_n\eta_{n-r}\\ &= (*,\cdots,*,k_{r+1},k_{r+2},\cdots,k_n),\end{aligned}$$

η'也是方程组(1)的一组解,它与 η 在 $x_{r+1},\cdots,x_n$ 处取相同的值,按前面的说明(ii),$\eta=\eta'$. 从而

$$\eta = k_{r+1}\eta_1 + k_{r+2}\eta_2 + \cdots + k_n\eta_{n-r}.$$

综合(i),(ii)即知 $\eta_1,\eta_2,\cdots,\eta_{n-r}$是方程组(1)的一个基础解系,它恰好包含 $n-r$ 个向量. ▌

这个定理的证明过程实质上指出了求方程组(1)的基础解系的具体办法:只要设法找到方程组的某一组自由未知量,由它们可以唯一确定其它未知量的值,那么,轮流让某个自由未知量取值1,其余取值0,确定出方程组的一个解向量,这样得到方程组的一组解向量,它们就是一个基础解系. 定理1又指明:不管这组自由未知量如何选法,其中未知量的数目都是 $n-r$ 个.

定理1有一个明显然而重要的推论.

推论 如果齐次线性方程组系数矩阵 A 的秩 r 等于未知量个数 n,则它只有零解. 而如果 $r<n$,它必有非零解. 或者说,齐次线性方程组有非零解的充分必要条件是其系数矩阵的秩 r 小于 n.

基础解系的求法

根据上一段落所说的，我们只要找到齐次线性方程组(1)的 $n-r$ 个自由未知量，就可以获得它的基础解系. 在§1中我们已经知道可以利用矩阵消元法来寻找方程组的自由未知量，所以矩阵消元法也就可以用来求齐次线性方程组的基础解系. 具体说，如果我们通过初等行变换把齐次线性方程组的系数矩阵(因为常数项为零，不必考虑增广矩阵)化为阶梯形，那么，阶梯形中阶梯的个数 r 即为系数矩阵的秩 $r(A)$. 把每个阶梯左角处对应的未知量保留在方程组左端，其余 $n-r$ 个未知量移到右端，此时，右端 $n-r$ 个未知量任取一组值，都可以唯一决定左端 r 个未知量的值. 所以，右端 $n-r$ 个未知量就是一组自由未知量(注意自由未知量的选取并不是唯一的).

例 1 求齐次线性方程组

$$\begin{cases} x_1+x_2-3x_4-x_5=0, \\ x_1-x_2+2x_3-x_4=0, \\ 4x_1-2x_2+6x_3+3x_4-4x_5=0, \\ 2x_1+4x_2-2x_3+4x_4-7x_5=0 \end{cases}$$

的一个基础解系.

解 在§1的例3中已把系数矩阵化为阶梯形

$$\begin{pmatrix} 1 & 1 & 0 & -3 & -1 \\ 1 & -1 & 2 & -1 & 0 \\ 4 & -2 & 6 & 3 & -4 \\ 2 & 4 & -2 & 4 & -7 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 0 & -3 & -1 \\ 0 & 2 & -2 & -2 & -1 \\ 0 & 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

现在 $r(A)=3$，基础解系中应有 $n-r=5-3=2$ 个向量. 写出阶梯形矩阵对应的方程组

$$\begin{cases} x_1+x_2-3x_4-x_5=0, \\ 2x_2-2x_3-2x_4-x_5=0, \\ 3x_4-x_5=0. \end{cases}$$

移项，得

$$\begin{cases} x_1 + x_2 - 3x_4 = x_5, \\ 2x_2 - 2x_4 = 2x_3 + x_5, \\ 3x_4 = x_5. \end{cases}$$

故 x_3, x_5 是自由未知量.

(i) 取 $x_3=1, x_5=0$，得一个解向量

$$\eta_1 = (-1,1,1,0,0).$$

(ii) 取 $x_3=0, x_5=1$，得另一个解向量

$$\eta_2 = \left(\frac{7}{6}, \frac{5}{6}, 0, \frac{1}{3}, 1\right).$$

η_1, η_2 即为方程组的一个基础解系，方程组的全部解可表示为

$$k_1\eta_1 + k_2\eta_2,$$

其中 k_1, k_2 为任意数.

例 2 求齐次线性方程组

$$\begin{cases} x_2 - x_3 + x_4 - x_5 = 0, \\ x_1 + x_3 + 2x_4 - x_5 = 0, \\ x_1 + x_2 + 3x_4 - 2x_5 = 0, \\ 2x_1 + 2x_2 + 6x_4 - 3x_5 = 0 \end{cases}$$

的一个基础解系.

解 先做矩阵消元法

$$\begin{pmatrix} 0 & 1 & -1 & 1 & -1 \\ 1 & 0 & 1 & 2 & -1 \\ 1 & 1 & 0 & 3 & -2 \\ 2 & 2 & 0 & 6 & -3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 1 & 2 & -1 \\ 0 & 1 & -1 & 1 & -1 \\ 1 & 1 & 0 & 3 & -2 \\ 2 & 2 & 0 & 6 & -3 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 0 & 1 & 2 & -1 \\ 0 & 1 & -1 & 1 & -1 \\ 0 & 1 & -1 & 1 & -1 \\ 0 & 2 & -2 & 2 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 1 & 2 & -1 \\ 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 0 & 1 & 2 & -1 \\ 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

$r(A)=3$，故基础解系中应包含 $n-r=5-3=2$ 个向量.写出阶梯形矩阵的对应方程组

$$\begin{cases} x_1 \quad + x_3 + 2x_4 - x_5 = 0, \\ \quad x_2 - x_3 + x_4 - x_5 = 0, \\ \qquad\qquad\qquad\qquad x_5 = 0. \end{cases}$$

移项，得

$$\begin{cases} x_1 \quad - x_5 = - x_3 - 2x_4, \\ \quad x_2 - x_5 = \quad x_3 - \quad x_4, \\ \qquad\qquad x_5 = 0. \end{cases}$$

x_3,x_4 为自由未知量.

(i) 取 $x_3=1$，$x_4=0$，得一个解向量

$$\eta_1 = (-1,1,1,0,0).$$

(ii) 取 $x_3=0$，$x_4=1$，得另一个解向量

$$\eta_2 = (-2,-1,0,1,0).$$

于是 η_1,η_2 为方程组的一个基础解系.方程组的全部解可表为

$$k_1\eta_1+k_2\eta_2,$$

其中 k_1,k_2 为任意数.

习　题　四

1. 求下列齐次线性方程组的一个基础解系

(1) $\begin{cases} x_1+x_2+x_3+x_4+x_5=0, \\ 3x_1+2x_2+x_3+x_4-3x_5=0, \\ \qquad x_2+2x_3+2x_4+6x_5=0, \\ 5x_1+4x_2+3x_3+3x_4-x_5=0. \end{cases}$

(2) $\begin{cases} x_1-2x_2+x_3+x_4-x_5=0, \\ 2x_1+x_2-x_3-x_4-x_5=0, \\ x_1+7x_2-5x_3-5x_4+5x_5=0, \\ 3x_1-x_2-2x_3+x_4-x_5=0. \end{cases}$

(3) $\begin{cases} x_1-2x_2+x_3-x_4+x_5=0, \\ 2x_1+x_2-x_3+2x_4-3x_5=0, \\ 3x_1-2x_2-x_3+x_4-2x_5=0, \\ 2x_1-5x_2+x_3-2x_4+2x_5=0. \end{cases}$

(4) $\begin{cases} 2x_1+6x_2-x_3+5x_4=0, \\ 3x_1-x_2+2x_3-7x_4=0, \\ 4x_1+x_2-3x_3+6x_4=0, \\ x_1-2x_2+4x_3-7x_4=0. \end{cases}$

(5) $\begin{cases} x_2-x_3+x_4=0, \\ -7x_2+3x_3+x_4=0, \\ x_1+3x_2-3x_4=0, \\ x_1-2x_2+3x_3-4x_4=0. \end{cases}$

(6) $\begin{cases} 2x_1+x_2-x_3-x_4+x_5=0, \\ x_1-x_2+x_3+x_4-2x_5=0, \\ 3x_1+3x_2-3x_3-3x_4+4x_5=0, \\ 4x_1+5x_2-5x_3-5x_4+7x_5=0. \end{cases}$

2. 证明：与基础解系等价的线性无关向量组也是基础解系.

3. 如果一个齐次线性方程组的系数矩阵 A 的秩为 r，证明：方程组的任意 $n-r$ 个线性无关的解向量都是它的一个基础解系.

4. 设

$$\alpha_i=(a_{i1},\ a_{i2},\ \cdots,\ a_{in})\quad (i=1,\ 2,\ \cdots,\ s);$$

$$\beta=(b_1,\ b_2,\ \cdots,\ b_n).$$

证明：如果齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{s1}x_1 + a_{s2}x_2 + \cdots + a_{sn}x_n = 0 \end{cases}$$

的解全是方程

$$b_1x_1 + b_2x_2 + \cdots + b_nx_n = 0$$

的解,那么,β 可以被 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性表示.

5. 给定两个齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0, \end{cases}$$

$$\begin{cases} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n = 0, \\ b_{21}x_1 + b_{22}x_2 + \cdots + b_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_{s1}x_1 + b_{s2}x_2 + \cdots + b_{sn}x_n = 0, \end{cases}$$

如果它们系数矩阵的秩都 $< n/2$,证明这两个方程组必有公共非零解.

6. 判断下面齐次线性方程组有无非零解

$$\begin{cases} \quad\;\; x_2 + x_3 + \cdots + x_n = 0, \\ x_1 \quad\;\; + x_3 + \cdots + x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_1 + x_2 + x_3 + \cdots + x_{n-1} \quad = 0 \end{cases}$$

(其中第 i 个方程缺 x_i).

§5 线性方程组的一般理论

现在来考察一般线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases} \tag{1}$$

其系数矩阵和增广矩阵分别是

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

$$\overline{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}.$$

在本节中,我们将从理论上对方程组(1)的解的情况进行详细的研究.

判 别 定 理

首先给出方程组(1)有解或无解的判别定理.

定理 2 线性方程组(1)有解的充分必要条件是其系数矩阵和增广矩阵的秩相等,即 $r(A)=r(\overline{A})$.

注意到 $r(A)$和 $r(\overline{A})$分别是将 $\overline{A}$ 用初等行变换化为阶梯形时前 n 列以及整个矩阵的阶梯数,因此,从§1关于矩阵消元法的初步理论讨论中,实际上已经得到了定理 2 的证明.在这里,再应用 m 维向量空间的一般理论来给出它的另一个证明.

定理 2 的证明 设 $r(A)=r$. $\overline{A}$ 的列向量组记为 $\alpha_1,\alpha_2,\cdots,\alpha_n$, β,其中前 n 个恰为 A 的列向量组.不妨设 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 为 A 的列向量组的一个极大线性无关部分组.下面分两个方面来证明定理 2.

(i) 必要性. 设方程组有解,证 $r(A)=r(\overline{A})$.

方程组有解，则 β 可被 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表示，从而可由 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性表示. 于是 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 也是 $\overline{A}$ 的列向量组的极大线性无关部分组，故 $r(\overline{A})=r(A)$.

(ii) 充分性. 设 $r(A)=r(\overline{A})$，证方程组(1)有解.

首先证 β 必能被 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性表示. 用反证法. 如果 β 不能被 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性表示，考察向量等式

$$k_1\alpha_1+\cdots+k_r\alpha_r+k\beta=0.$$

上式中必有 $k=0$. 因若 $k\neq0$，则

$$\beta=-\frac{k_1}{k}\alpha_1-\cdots-\frac{k_r}{k}\alpha_r.$$

这与 β 不能被 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性表示之假设矛盾. 而 $k=0$ 时，有

$$k_1\alpha_1+\cdots+k_r\alpha_r=0.$$

由 $\alpha_1,\cdots,\alpha_r$ 线性无关，推知 $k_1=\cdots=k_r=0$. 这一事实说明向量组 $\alpha_1,\alpha_2,\cdots,\alpha_r,\beta$ 线性无关，故 $\alpha_1,\alpha_2,\cdots,\alpha_r,\beta$ 是 $\overline{A}$ 的列向量组的一个极大线性无关部分组. 于是 $r(\overline{A})=r+1\neq r(A)$，与假设矛盾. 因此，$\beta$ 必定能被 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性表示，因而它也就能被 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表示(取 $\alpha_{r+1},\cdots,\alpha_n$ 的系数为零即可). 这说明方程组(1)有解. ▌

解的结构

下面再来讨论方程组(1)有解的情况下解的结构问题.

把方程组(1)的常数项通通换成 0，得到与之对应的一个齐次线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0,\\a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0,\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0.\end{cases}\tag{2}$$

方程组(2)称为方程组(1)的**导出方程组**. 方程组(1)与(2)的解之间有如下关系：

(i) 设

$$\gamma_1 = (c_1, c_2, \cdots, c_n), \gamma_2 = (d_1, d_2, \cdots, d_n)$$

是方程组(1)的两个解向量，则 $\eta=\gamma_1-\gamma_2$ 是方程组(2)的一个解向量；

(ii) 设 $\gamma_0=(c_1, c_2, \cdots, c_n)$是方程组(1)的一个解向量，而 $\eta=(k_1, k_2, \cdots, k_n)$是方程组(2)的一个解向量，则 $\gamma=\gamma_0+\eta$ 是方程组(1)的一个解向量.

这两条性质只要直接验证就可以了. 例如，对性质(i)，有

$$\left.\begin{aligned}\sum_{j=1}^{n} a_{ij}c_j = b_i\\ \sum_{j=1}^{n} a_{ij}d_j = b_i\end{aligned}\right. \quad (i = 1, 2, \cdots, m).$$

两式相减，得

$$\sum_{j=1}^{n} a_{ij}c_j - \sum_{j=1}^{n} a_{ij}d_j = \sum_{j=1}^{n} a_{ij}(c_j - d_j) = 0.$$

这表示 $\eta=\gamma_1-\gamma_2$ 是方程组(2)的解. 性质(ii)请读者自行验证.

从上面两条简单的性质可以知道，如果给定方程组(1)的某一个解向量 γ_0，那么，对于方程组(1)的任一个解向量 γ，$\gamma-\gamma_0=\eta$ 是方程组(2)的一个解向量，故 γ 可表示为

$$\gamma = \gamma_0 + \eta.$$

反之，γ_0 加上方程组(2)的任一解向量 η 就得到方程组(1)的一个解向量. 如果设方程组(2)的一个基础解系为 $\eta_1, \eta_2, \cdots, \eta_{n-r}$，则方程组(2)的全部解可用下式表出

$$k_1\eta_1 + k_2\eta_2 + \cdots + k_{n-r}\eta_{n-r}.$$

从而方程组(1)的全部解可表示为

$$\gamma = \gamma_0 + k_1\eta_1 + k_2\eta_2 + \cdots + k_{n-r}\eta_{n-r}, \tag{3}$$

其中 $k_1, k_2, \cdots, k_{n-r}$可取任意数.

定理 3 在 $r(A)=r(\overline{A})$的条件下，有

(i) 如果 $r(A)=n$，则方程组(1)有唯一解；

(ii) 如果 $r(A)<n$，则方程组(1)有无穷多组解，其全部解可由某一特殊解 γ_0 和它的导出方程组(2)的一个基础解系用(3)式表示.

证 (i) 如果 $r(A)=n$，则 A 的列向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关，$\overline{A}$ 的最后一列 β 被 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表示的表法是唯一的(命题 4.1)，故方程组(1)的解唯一.

(ii) 如果 $r(A)<n$，则齐次线性方程组(2)有基础解系，其正确性已在前面讨论中获知. ▌

至此，本章开头所提出的关于线性方程组解的三个理论问题全部获得解决.

根据定理 3，求方程组(1)的全部解可由下列两部分工作组成：

(i) 求方程组(1)的某一特解 γ_0；

(ii) 求导出方程组(2)的一个基础解系.

最后，代入公式(3)即得方程组(1)的全部解. 而这两部分工作可用矩阵消元法同时进行.

例 1 求方程组

$$\begin{cases} x_1 - x_2 \qquad + x_4 - x_5 = 1, \\ 2x_1 \qquad + x_3 \qquad - x_5 = 2, \\ 3x_1 - x_2 - x_3 - x_4 - x_5 = 0 \end{cases}$$

的全部解.

解 写出方程组的增广矩阵，对它做初等行变换化为阶梯形

$$\overline{A} = \begin{pmatrix} 1 & -1 & 0 & 1 & -1 & 1 \\ 2 & 0 & 1 & 0 & -1 & 2 \\ 3 & -1 & -1 & -1 & -1 & 0 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & -1 & 0 & 1 & -1 & 1 \\ 0 & 2 & 1 & -2 & 1 & 0 \\ 0 & 2 & -1 & -4 & 2 & -3 \end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1 & -1 & 0 & 1 & -1 & 1\\ 0 & 2 & 1 & -2 & 1 & 0\\ 0 & 0 & 2 & 2 & -1 & 3\end{pmatrix}.$$

写出对应的方程组

$$\begin{cases}x_1-x_2+x_4-x_5=1,\\ 2x_2+x_3-2x_4+x_5=0,\\ 2x_3+2x_4-x_5=3.\end{cases}$$

这方程组与原方程组同解,因此,只要求它的全部解就可以了.移项,得

$$\begin{cases}x_1-x_2=1-x_4+x_5,\\ 2x_2+x_3=2x_4-x_5,\\ 2x_3=3-2x_4+x_5.\end{cases}$$

(i) 先求一个特解 γ_0. 这只要取 $x_4=x_5=0$ 即可. 有

$$\gamma_0=\left(\frac{1}{4},-\frac{3}{4},\frac{3}{2},0,0\right).$$

(ii) 再求它的导出方程组的基础解系. 这只要把方程组的常数项换成零,得

$$\begin{cases}x_1-x_2=-x_4+x_5,\\ 2x_2+x_3=2x_4-x_5,\\ 2x_3=-2x_4+x_5.\end{cases}$$

x_4,x_5 为自由未知量.

$$x_4=1,\ x_5=0:\quad \eta_1=\left(\frac{1}{2},\frac{3}{2},-1,1,0\right);$$

$$x_4=0,\ x_5=1:\quad \eta_2=\left(\frac{1}{4},-\frac{3}{4},\frac{1}{2},0,1\right).$$

故原方程组的全部解为

$$\gamma=\left(\frac{1}{4},-\frac{3}{4},\frac{3}{2},0,0\right)+k_1\left(\frac{1}{2},\frac{3}{2},-1,1,0\right)$$
$$+k_2\left(\frac{1}{4},-\frac{3}{4},\frac{1}{2},0,1\right),$$

其中 k_1,k_2 可取任意数.

数域的概念

读者可以看到,本章中所涉及到的数的运算仅有加、减、乘、除四种. 根据 § 1 最后所作的说明,我们在讨论本章的各项课题时,实际上都可以限制在数的某些范围内进行.

定义 由某些数组成的一个集合 K,如果其中至少包含两个不同的数,且它对数的加、减、乘、除四种运算是封闭的,即 K 中任意两个数相加、相减、相乘、相除(但不能以 0 作除数)所得的数仍然属于 K,则 K 就称为一个**数域**.

例 2 全体有理数所成的集合是一数域. 这个数域通常用字母 $\mathscr{Q}$ 表示,称为有理数域.

例 3 全体实数所成的集合是一数域. 这个数域通常用字母 $\mathscr{R}$ 表示,称为实数域.

例 4 全体复数所成的集合是一数域. 这个数域通常用字母 $\mathscr{C}$ 表示,称为复数域.

例 5 全体形如 $a+bi$ 的数(其中 a,b 是有理数)的集合是一数域,因为不难验证它对加、减、乘、除是封闭的. 由于这里 a,b 仅限于取有理数值,所以这个数域只是复数域的一部分.

全体整数所成的集合不是数域,因为它对除法不封闭.

有了数域的概念之后,前面的讨论都可以限制在某个数域内来进行. 例如,以数域 K 的元素作系数和常数项的线性方程组,我们称之为**数域 K 上的线性方程组**. 对这样的方程组的一切讨论,包括它的自由未知量的取值等等,都可以限制在数域 K 内来进行. 又比如,讨论以数域 K 的元素作分量的全体 n 维向量所成的集合,这个集合关于向量加法以及与 K 中元素的数乘运算是封闭的,因而,对它的讨论也可以限制在数域 K 内进行,我们就称之为**数域 K 上的 n 维向量空间**. 例如实数域 $\mathscr{R}$ 上的 n 维向量空间;复数域 $\mathscr{C}$ 上的 n 维向量空间,等等.

把问题限制在某个数域内来讨论，这对某些问题是重要的．本书后面的章节中将要涉及这一点，因此在这里对数域的概念预先作一介绍，以使读者有所准备．

习　题　五

1． 求下列线性方程组的一个特解 γ_0 和导出方程组的一个基础解系，然后用它们表出方程组的全部解．

(1) $$\begin{cases} 2x_1-2x_2+x_3-x_4+x_5=1, \\ x_1+2x_2-x_3+x_4-2x_5=1, \\ 4x_1-10x_2+5x_3-5x_4+7x_5=1, \\ 2x_1-14x_2+7x_3-7x_4+11x_5=-1. \end{cases}$$

(2) $$\begin{cases} 9x_1+12x_2+17x_4=4, \\ 6x_1+3x_2+3x_3+8x_4=4, \\ 3x_1+6x_2+6x_3+13x_4=2, \\ x_1+6x_2+2x_3+9x_4=0, \\ 6x_1+5x_2-x_3+7x_4=3. \end{cases}$$

2． 证明：如果 $\eta_1,\eta_2,\cdots,\eta_t$ 是一个线性方程组的解，那么 $k_1\eta_1+k_2\eta_2+\cdots+k_t\eta_t$（其中 $k_1+k_2+\cdots+k_t=1$）也是一个解．

3． 在什么条件下 n 个平面

$$A_ix+B_iy+C_iz+D_i=0 \quad (i=1,2,\cdots,n)$$

通过同一点？在什么条件下它们通过同一直线？

4． 给定线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n. \end{cases}$$

令

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},\ B=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \\ b_1 & b_2 & \cdots & b_n & 0 \end{pmatrix}.$$

证明：若 $r(A)=r(B)$，则方程组有解.

5. 给定方程组

$$\begin{cases} ax_1+ax_2+\cdots+ax_{n-1}+bx_n=b_1, \\ ax_1+ax_2+\cdots+bx_{n-1}+ax_n=b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ bx_1+ax_2+\cdots+ax_{n-1}+ax_n=b_n, \end{cases}$$

试对 a,b 及 $b_1,b_2,\cdots,b_n$ 讨论方程组何时有解？有解时有多少组解？

6. 设 γ_0 是线性方程组的一个特解，$\eta_1,\eta_2,\cdots,\eta_s$ 是其导出方程组的一个基础解系. 令

$$\gamma_1=\gamma_0+\eta_1,\ \gamma_2=\gamma_0+\eta_2,\ \cdots,\ \gamma_s=\gamma_0+\eta_s,$$

证明：线性方程组的任一解 γ 可表成

$$\gamma=k_0\gamma_0+k_1\gamma_1+\cdots+k_s\gamma_s,$$

其中 $k_0+k_1+\cdots+k_s=1$.

7. 证明：全体形如 $a+b\sqrt{3}$（a,b 取有理数）的数所组成的集合是一个数域.

8. 证明：一个数域 K 内必包含 0,1 两个数. 又若 a 属于数域 K，$a\neq0$，则 $1/a$ 也属于数域 K.

9. 证明：有理数域 $\mathcal{Q}$ 包含在任意一个数域 K 之中.

10. 举出一个具体的数域，在其中开方运算并不总能进行（即 K 内有一个数 a，使方程 $x^2-a=0$ 在 K 内没有根）.

第二章 矩阵代数

在第一章里，为了讨论线性方程组，我们引进了矩阵的概念. 这个概念在线性方程组的理论中发挥了重要的作用. 但是在那里，矩阵还仅仅是一个“表格”，其内容还是比较贫乏的，这就使它所起的作用受到很大的限制. 在这一章里，我们将要在矩阵之间引进某些运算，从而充实了矩阵概念的内容，使它丰富起来. 这样，不但对我们进一步研究线性方程组是有用的，而且也大大扩充了矩阵的应用领域，使它在数学以至自然科学、工程技术的许多问题中都成了一个有力的工具.

本章的主要内容是介绍矩阵的加法、数乘和乘法三种运算以及与这三种运算相关联的一些命题和定理. 本章的内容在矩阵理论和实际应用中都具有基本的意义.

§1 矩阵的运算

在这一节里，我们给出矩阵三种基本运算的定义和一些基本性质.

矩阵的加法和数乘

定义 给定数域 K 上两个 $m\times n$ 矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},\quad B=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix},$$

它们的加法定义为：

$$A+B=\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{pmatrix}.$$

对任意 $k\in K$，k 与 A 的数乘定义为：

$$kA=\begin{pmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & & \vdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{pmatrix}.$$

读者应当注意：只有既同行数又同列数的矩阵才能相加. 另外，两个矩阵相等是指它们的元素完全相同.

不难看出，矩阵的加法和数乘本质上与向量的加法和数乘是一样的. 实际上，一个 n 维向量可以当作一个 $1\times n$ 矩阵(如果它写成横排方式)或 $n\times 1$ 矩阵(如果它写成竖列方式). 由于这个原因，第一章命题 2.1 所列举的向量加法、数乘运算的八条性质对矩阵的加法和数乘运算也完全适用，只要把那里的向量 α,β,γ 换成矩阵 A,B,C 就可以了. 在这里，零矩阵(元素全为零的矩阵)代替了零向量的地位，而 A 的"负矩阵"

$$-A=\begin{pmatrix} -a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & -a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{m1} & -a_{m2} & \cdots & -a_{mn} \end{pmatrix}$$

代替了那里的"负向量"的地位. 同样，把 $A+(-B)$ 写成 $A-B$，称为矩阵的减法运算.

在这里，我们所考察的是其元素属于某一数域 K 的 $m\times n$ 矩阵，而作数乘的数也仅限于数域 K 的数. 这样，就把讨论的范围限制在数域 K 内. 由于在本章中我们所涉及的数的运算也仅限于加、减、乘、除四种，所以始终不会越出数域 K 的范围. 因此，我们可以说，本章所讨论的，是数域 K 上的矩阵的基本代数运算.

矩阵的乘法

现在我们来引进矩阵的第三种运算：乘法. 这是本章的核心内容.

矩阵的乘法是从线性方程组的研究中自然产生出来的. 给定线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m. \end{cases}$$

在第一章里，我们用增广矩阵代表这个方程组. 现在，我们要把这个方程组用另一种形式表示出来. 为此，考察如下两个矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

这就是方程组的系数矩阵 A 和未知量所组成的 $n\times 1$ 矩阵 X. 我们规定 A 与 X 的“乘法”如下：

$$AX=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$=\begin{pmatrix} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n \end{pmatrix}.$$

其法则是：左边矩阵的行和右边矩阵的列对应元素相乘再相加. 乘得的结果是一个 $m\times 1$ 矩阵(上面用虚线在矩阵中标出左边的行

与右边的列对应相乘的位置关系). 显然,乘积矩阵自上而下的第 $1,2,\cdots,m$ 个元素恰好是方程组的第 $1,2,\cdots,m$ 个方程的左端. 如再引入如下 $m\times 1$ 矩阵

$$B=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

那么,方程组可用如下矩阵方程表示

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

或用抽象的记号,写成

$$AX=B.$$

这样,方程组的系数(用矩阵 A 代表)和未知量(用矩阵 X 代表)及常数项(用矩阵 B 代表)这三者之间的互相制约关系就借助于上面引入的矩阵乘法运算清楚地表达出来了.

现在线性方程组一共有了三种新表现方式:(i) 用其增广矩阵 $\overline{A}$ 表示;(ii) 用向量方程表示;(iii) 用矩阵方程表示. 这三种表现形式各有各的用处,读者都应当熟悉. 现在将线性方程组表成 $AX=B$,它在形式上很像初等代数中的一元一次方程式 $ax=b$. 一元一次方程式的解为 $x=a^{-1}b(a\neq 0)$. 下面将会看到,在一定条件下,矩阵方程 $AX=B$ 也可以类似地解出.

例 1 给定线性方程组

$$\begin{cases} x_1-x_2 \qquad\quad +x_4-x_5=1, \\ 2x_1 \qquad +x_3 \qquad\quad -x_5=2, \\ 3x_1-x_2-x_3-x_4-x_5=0, \end{cases}$$

那么,可以把它表示成如下矩阵方程式

$$\begin{pmatrix}1 & -1 & 0 & 1 & -1\\ 2 & 0 & 1 & 0 & -1\\ 3 & -1 & -1 & -1 & -1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\\ x_4\\ x_5\end{pmatrix}=\begin{pmatrix}1\\ 2\\ 0\end{pmatrix}.$$

下面,我们再把矩阵乘积 AX 中的 X 换成由普通数组成的矩阵,例如,令

$$B=\begin{pmatrix}b_1\\ b_2\\ \vdots\\ b_n\end{pmatrix}$$

(这里的 B 是 $n\times 1$ 矩阵,不是上面方程组常数项组成的 $m\times 1$ 矩阵),那么,应当有

$$AB=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix}\begin{pmatrix}b_1\\ b_2\\ \vdots\\ b_n\end{pmatrix}$$

$$=\begin{pmatrix}a_{11}b_1+a_{12}b_2+\cdots+a_{1n}b_n\\ a_{21}b_1+a_{22}b_2+\cdots+a_{2n}b_n\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ a_{m1}b_1+a_{m2}b_2+\cdots+a_{mn}b_n\end{pmatrix}.$$

现在乘积矩阵是一个 $m\times 1$ 的数值矩阵.

例 2 作下面的矩阵乘法

$$\begin{pmatrix}-1 & 0 & 1 & 0\\ 2 & 1 & 0 & 1\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1\\ -1\\ 0\\ 1\end{pmatrix}=\begin{pmatrix}-1+0+0+0\\ 2-1+0+1\\ 0+0+0+0\\ 0-1+0+0\\ 1+0+0+1\end{pmatrix}=\begin{pmatrix}-1\\ 2\\ 0\\ -1\\ 2\end{pmatrix}.$$

这里所定义的 $m\times n$ 矩阵(在左边)和 $n\times 1$ 矩阵(在右边)的乘法法则是矩阵乘法运算的基础,把它搞清楚了,一般矩阵的乘法也就不难明白了.在上面的定义中包含了如下几个要点:

(1) 左边矩阵 A 的列数必须等于右边矩阵 B 的行数才能相乘;

(2) 乘积矩阵的行数等于左边矩阵的行数,其列数则等于右边矩阵的列数(现在是 1);

(3) 乘法的法则是左边矩阵的行和右边矩阵的列的对应元素相乘再相加;

(4) 如果把左边矩阵 A 的列向量组记为 $\alpha_1,\alpha_2,\cdots,\alpha_n$(都是 m 维向量),那么,乘积矩阵也可以看作一个 m 维向量,它正好等于

$$b_1\alpha_1+b_2\alpha_2+\cdots+b_n\alpha_n.$$

也就是说,上面定义的矩阵乘法本质上是用右边矩阵 B 的元素作为系数去作左边矩阵 A 的列向量组的线性组合.如果把 B 再改为 X,那么 AX 就是

$$x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n.$$

这正好与第一章 § 2 中将线性方程组表成向量方程的结果一样.

如果用和号表示上面引进的矩阵运算,就是

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}=\begin{bmatrix} \sum\limits_{k=1}^{n} a_{1k}b_k \\ \sum\limits_{k=1}^{n} a_{2k}b_k \\ \vdots \\ \sum\limits_{k=1}^{n} a_{mk}b_k \end{bmatrix}.$$

在一般情况下,当右边的矩阵不止一列,而是有 s 列时,乘法的法则是:把左边矩阵 A 分别乘右边矩阵的每一列(按上面所述法则),然后把所得的 s 个列矩阵依次排列,即得乘积矩阵.我们把

它表述成如下的

定义 给定数域 K 上的 $m\times n$ 矩阵 A 和 $n\times s$ 矩阵 B

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad B=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1s} \\ b_{21} & b_{22} & \cdots & b_{2s} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{ns} \end{pmatrix},$$

定义 A 与 B 的乘法如下：

$$AB=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1s} \\ b_{21} & b_{22} & \cdots & b_{2s} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{ns} \end{pmatrix}$$

$$=\begin{pmatrix} \sum\limits_{k=1}^{n} a_{1k}b_{k1} & \sum\limits_{k=1}^{n} a_{1k}b_{k2} & \cdots & \sum\limits_{k=1}^{n} a_{1k}b_{ks} \\ \sum\limits_{k=1}^{n} a_{2k}b_{k1} & \sum\limits_{k=1}^{n} a_{2k}b_{k2} & \cdots & \sum\limits_{k=1}^{n} a_{2k}b_{ks} \\ \vdots & \vdots & & \vdots \\ \sum\limits_{k=1}^{n} a_{mk}b_{k1} & \sum\limits_{k=1}^{n} a_{mk}b_{k2} & \cdots & \sum\limits_{k=1}^{n} a_{mk}b_{ks} \end{pmatrix}=C.$$

这个定义中所包含的矩阵乘法运算的要点和前面对 B 是 $n\times 1$矩阵情况下论述的一样，只是对(ii)要作一点补充：乘积矩阵 C 的列数现在等于右边的矩阵 B 的列数 s. 另外，还要强调一下：C 的第 i 个列向量($i=1,2,\cdots,s$)是用 B 的第 i 列元素作为系数去作 A 的列向量组的线性组合得来的.

例 3 作矩阵乘法

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ -1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}=\begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 2 & 2 & 1 \\ -1 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 \end{pmatrix}.$$

例 4 给定矩阵

$$A=\begin{pmatrix}0&1&0\\0&0&1\\0&0&1\end{pmatrix},$$

$$B=\begin{pmatrix}0&0&1\\0&0&0\\0&0&0\end{pmatrix},$$

则有

$$AB=\begin{pmatrix}0&1&0\\0&0&1\\0&0&1\end{pmatrix}\begin{pmatrix}0&0&1\\0&0&0\\0&0&0\end{pmatrix}=\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}=0,$$

$$BA=\begin{pmatrix}0&0&1\\0&0&0\\0&0&0\end{pmatrix}\begin{pmatrix}0&1&0\\0&0&1\\0&0&1\end{pmatrix}\begin{pmatrix}0&0&1\\0&0&0\\0&0&0\end{pmatrix}\neq 0.$$

从例 4 可以看出，矩阵乘法和数的乘法有两个根本性的区别：

(i) 矩阵乘法一般是不可交换的，即在一般情况下，$AB\neq BA$. 实际上，AB 有意义时，BA 不一定有意义，因为 B 的列数不一定等于 A 的行数. 而且，即使有意义，两者也不一定相等.

(ii) 两个非零矩阵相乘有可能变成零矩阵. 因而，由 $AB=0$ 并不能推出 $A=0$ 或 $B=0$. 随之而来的是：由 $AB=AC$，且 $A\neq 0$，并不能推出 $B=C$(即消去律不成立).

矩阵乘法的基本性质

矩阵乘法满足如下运算规律：

(1) 结合律：$A(BC)=(AB)C$；

(2) 分配律：$(A+B)C=AC+BC$，

$$A(B+C)=AB+AC;$$

(3) 对任一数 k，有 $k(AB)=(kA)B=A(kB)$.

我们来证明结合律：设

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},\quad B=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1l} \\ b_{21} & b_{22} & \cdots & b_{2l} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nl} \end{pmatrix},$$

$$C=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1s} \\ c_{21} & c_{22} & \cdots & c_{2s} \\ \vdots & \vdots & & \vdots \\ c_{l1} & c_{l2} & \cdots & c_{ls} \end{pmatrix}.$$

显然，此时二重乘积 $A(BC)$ 与 $(AB)C$ 都有意义. 为了简便，上面三个矩阵简记为：$A=(a_{ij})_{m\times n}$；$B=(b_{ij})_{n\times l}$，$C=(c_{ij})_{l\times s}$.

此时，矩阵 AB 的 i 行 t 列元素为 $\sum\limits_{k=1}^{n}a_{ik}b_{kt}$，故

$$AB=\left(\sum_{k=1}^{n}a_{ik}b_{kt}\right)_{m\times l}.$$

从而 $(AB)C$ 的 i 行 j 列元素为

$$\begin{aligned}\sum_{t=1}^{l}\left(\sum_{k=1}^{n}a_{ik}b_{kt}\right)c_{tj} &= \sum_{k=1}^{n}\sum_{t=1}^{l}a_{ik}b_{kt}c_{tj} \\ &= \sum_{k=1}^{n}a_{ik}\left(\sum_{t=1}^{l}b_{kt}c_{tj}\right). \end{aligned} \tag{1}$$

再考察矩阵 BC，其 k 行 j 列元素为 $\sum\limits_{t=1}^{l}b_{kt}c_{tj}$，故

$$BC=\left(\sum_{t=1}^{l}b_{kt}c_{tj}\right)_{n\times s}.$$

从而 $A(BC)$ 的 i 行 j 列元素为

$$\sum_{k=1}^{n}a_{ik}\left(\sum_{t=1}^{l}b_{kt}c_{tj}\right). \tag{2}$$

比较(1)与(2)，即得 $(AB)C=A(BC)$.

性质(2)与(3)请读者自行证明.

下面介绍几类重要的矩阵.

1. n 阶方阵.

一个 n 行 n 列矩阵 $A=(a_{ij})_{n\times n}$称为一个 n **阶方阵**. 一个 n 阶方阵从左上角到右下角元素间的连线(由 $a_{11},a_{22},\cdots,a_{nn}$组成)称为它的**主对角线**.

2. n 阶单位矩阵.

命

$$E_n=\begin{pmatrix}1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1\end{pmatrix}_{n\times n}$$

(即主对角线上元素全是 1,主对角线外元素全是 0). E_n 称为 n 阶**单位矩阵**. 当其阶数从上下文看很明确时,也可省略其下角标,简记为 E.

3. 对角矩阵和数量矩阵.

称下列矩阵

$$D_n=\begin{pmatrix}d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n\end{pmatrix}_{n\times n}$$

为 n 阶**对角矩阵**(矩阵中空白处元素为零,下同). 而当 $d_1=d_2=\cdots=d_n=d$ 时,称为**数量矩阵**. 显然,数量矩阵可写成 dE 形式. 对角矩阵在矩阵乘法中占有特殊的地位,原因是它有如下性质:

(1) 用一个 m 阶对角矩阵 D_m 左乘一个 $m\times n$ 矩阵 $A=(a_{ij})$,所得的结果相当于分别用 $d_1,d_2,\cdots,d_m$ 乘 A 的第 $1,2,\cdots,m$ 行;

(2) 用一个 n 阶对角矩阵 D_n 右乘 A,所得的结果相当于分别用 $d_1,d_2,\cdots,d_n$ 乘 A 的第 $1,2,\cdots,n$ 列.

具体写出来,就是

$$D_mA=\begin{pmatrix}d_1a_{11} & d_1a_{12} & \cdots & d_1a_{1n}\\ d_2a_{21} & d_2a_{22} & \cdots & d_2a_{2n}\\ \vdots & \vdots & & \vdots\\ d_ma_{m1} & d_ma_{m2} & \cdots & d_ma_{mn}\end{pmatrix};$$

$$AD_n=\begin{pmatrix} d_1a_{11} & d_2a_{12} & \cdots & d_na_{1n} \\ d_1a_{21} & d_2a_{22} & \cdots & d_na_{2n} \\ \vdots & \vdots & & \vdots \\ d_1a_{m1} & d_2a_{m2} & \cdots & d_na_{mn} \end{pmatrix}.$$

特别地，取 $D_n=E_n$ 为 n 阶单位矩阵，得

$$E_mA=AE_n=A.$$

如果 A 是 n 阶方阵，E 是 n 阶单位矩阵，有

$$EA=AE=A.$$

即 E 在矩阵乘法中的地位相当于数 1 在数的乘法运算中的地位.

如果 A 是一个 n 阶方阵，我们定义

$$A^0=E,\quad A^k=\underbrace{AA\cdots A}_{k\text{个}}.$$

这样，n 阶方阵的非负整数次幂都有了意义. 又若

$$f(\lambda)=a_0\lambda^m+a_1\lambda^{m-1}+\cdots+a_{m-1}\lambda+a_m$$

是 λ 的一个多项式，则定义

$$f(A)=a_0A^m+a_1A^{m-1}+\cdots+a_{m-1}A+a_mE,$$

称为矩阵 A 的多项式.

最后，我们来讨论第一章§3 中所引进的转置矩阵的一些基本性质. 我们有

(1) $(A')'=A$;

(2) $(A+B)'=A'+B'$;

(3) $(AB)'=B'A'$;

(4) $r(A)=r(A')$.

性质(1)，(2)，(4)都是显然的. 我们来证明性质(3)：

设 $A=(a_{ij})_{m\times n}$，$B=(b_{ij})_{n\times l}$. 矩阵 AB 的 j 行 i 列元素为 $\sum_{k=1}^{n}a_{jk}b_{ki}$，它就是$(AB)'$的 i 行 j 列元素. 另一方面，矩阵 B' 的 i 行 k 列元素为 B 的 k 行 i 列元素 b_{ki}；矩阵 A' 的 k 行 j 列元素为 A 的 j 行 k 列元素 a_{jk}. 故 $B'A'$ 的 i 行 j 列元素为

$$\sum_{k=1}^{n} b_{ki} a_{jk} = \sum_{k=1}^{n} a_{jk} b_{ki},$$

它恰等于$(AB)'$的i行j列元素. 于是$(AB)' = B'A'$. 读者应注意后面乘积的次序颠倒了. 一般讲, $(AB)' \neq A'B'$.

习 题 一

1. 进行下列矩阵运算:

(1) $$5\begin{pmatrix} 3 & -2 & -1 \\ 2 & 0 & 1 \\ -1 & 1 & 2 \\ 0 & 1 & 0 \end{pmatrix} - \begin{pmatrix} -1 & -1 & -1 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \\ 2 & 2 & 2 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 2 \\ 1 & -1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

(2) $$\begin{pmatrix} -1 & 3 & 2 \\ 4 & -1 & 0 \\ 0 & 2 & -1 \\ 1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}; \quad \begin{pmatrix} \mathrm{i} & -\mathrm{i} \\ 1 & 2+\mathrm{i} \\ 0 & -1 \\ 2-\mathrm{i} & \mathrm{i} \end{pmatrix}\begin{pmatrix} \mathrm{i} \\ -\mathrm{i} \end{pmatrix};$$

$$\begin{pmatrix} -3 & 1 & 0 & 2 & 1 \\ -1 & 0 & 1 & -1 & 1 \\ 0 & 2 & 0 & 1 & -1 \\ -1 & 1 & 0 & -2 & 1 \end{pmatrix}\begin{pmatrix} -1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix};$$

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 2 \\ 0 & 0 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 0 \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & 1 & -1 & 1 \end{pmatrix}.$$

(3) $\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}$;

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix}.$$

2. 设

(1) $A=\begin{pmatrix} 3 & 1 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix}$, $B=\begin{pmatrix} 1 & 1 & -1 \\ 2 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$;

(2) $A=\begin{pmatrix} a & b & c \\ c & b & a \\ 1 & 1 & 1 \end{pmatrix}$, $B=\begin{pmatrix} 1 & a & c \\ 1 & b & b \\ 1 & c & a \end{pmatrix}$.

计算 $AB, AB-BA, (AB)', A'B'$.

3. 计算:

(1) $\begin{pmatrix} 2 & 1 & 1 \\ 3 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}^2$; (2) $\begin{pmatrix} 3 & 2 \\ -4 & -2 \end{pmatrix}^5$;

(3) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n$; (4) $\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}^n$;

(5) $(2 \quad 3 \quad -1)\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$; (6) $\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}(2 \quad 3 \quad -1)$;

(7) $(x \quad y \quad 1)\begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{12} & a_{22} & b_2 \\ b_1 & b_2 & c \end{pmatrix}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$;

(8) $\begin{pmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix}^n$；(9) $\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}^n$.

4. 给定 n 阶方阵

$$J=\begin{pmatrix} 0 & 1 & & \\ & & \ddots & \\ & 0 & & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}_{n\times n},$$

证明：当 $k\geqslant n$ 时，$J^k=0$（矩阵中空白处元素为零）.

5. 设 $f(\lambda)=\lambda^2-\lambda-1$，而

$$A=\begin{pmatrix} 2 & 1 & 1 \\ 3 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix},$$

求 $f(A)$.

6. 如果两个 n 阶方阵满足 $AB=BA$，则称**可交换**. 求所有与

$$J=\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

可交换的 3 阶方阵.

7. 设给定对角矩阵

$$A=\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},\quad \lambda_i\neq\lambda_j\quad (i\neq j),$$

证明：与 A 可交换的矩阵都是对角矩阵.

8. 证明：如果 n 阶方阵 A 与所有 n 阶方阵都可交换，则 A 必是一个数量矩阵：$A=\lambda E$.

9. 如果一个 n 阶方阵 A 满足 $A'=A$，则称为**对称矩阵**. 证明：若 A 是一个 n 阶方阵，则 $A+A'$，AA'，$A'A$ 都是对称矩阵.

10. 设 A,B 都是对称矩阵，证明：AB 是对称矩阵当且仅当 A,B 可交换.

11. 如果一个 n 阶方阵 A 满足 $A'=-A$，则称为**反对称矩阵**. 证明：

(i) 若 A 是反对称矩阵，则 A 主对角线上的元素全为零；

(ii) 对任一 n 阶方阵 A，$A-A'$ 为反对称矩阵；

(iii) 如果 A,B 是两个 n 阶反对称矩阵，k,l 是任意两个数，则 $kA+lB$，$AB-BA$ 都是反对称矩阵.

12. 证明：任一 n 阶方阵都可以表示成一个对称矩阵和一个反对称矩阵之和.

13. 设 A 是一个 n 阶方阵，$A\neq 0$. 证明：存在一个非零的 n 阶方阵 B，使 $AB=0$ 的充分必要条件是

$$r(A)<n.$$

14. 举例说明：对一个非零的 n 阶方阵 A，当 $r(A)<n$ 时，由 $AB=AC$，不一定有

$$B=C.$$

15. 设 A,B 均为 n 阶方阵，且 $AB=0$，证明：

$$r(A)+r(B)\leqslant n.$$

(提示：利用第一章定理 1.)

16. 设 A 是 n 阶方阵. 证明：

(1) 若 $A^2=E$，则

$$r(A+E)+r(A-E)=n;$$

(2) 若 $A^2=A$，则

$$r(A)+r(A-E)=n.$$

17. 设 A 为 $m\times n$ 矩阵，B 为 $n\times s$ 矩阵，证明：

$$r(AB)\geqslant r(A)+r(B)-n.$$

(提示：设 $AB=C$，考察以 A 为系数矩阵，以 C 的每个列向量为常数项的一些线性方程组，利用第一章定理 3 和定理 1.)

18. 下列 n 阶方阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & 0 & & a_{nn} \end{pmatrix}$$

称为 n 阶**上三角矩阵**. 而

$$A=\begin{pmatrix} a_{11} & & & \\ a_{21} & a_{22} & & 0 \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

称为 n 阶**下三角矩阵**. 证明:两个上(下)三角矩阵的乘积仍为上(下)三角矩阵.

19. 设 A,B 为 n 阶方阵,且 $AB=BA$,证明:

$$(A+B)^n=\sum_{k=0}^{n}C_n^kA^kB^{n-k}.$$

20. 举例说明:当 A,B 为 n 阶方阵,而 $AB\neq BA$ 时

$$(A+B)(A-B)\neq A^2-B^2.$$

21. 求

$$\begin{pmatrix} \lambda_0 & 1 & & \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0 \end{pmatrix}_{n\times m}^k$$

(矩阵中空白处元素为零).

§2 初 等 矩 阵

在第一章§3中,我们引进了矩阵的初等变换的概念,这个概念对我们研究矩阵是一个有用的工具. 现在我们来指出:矩阵的初等行(列)变换可以用矩阵乘法表示出来.

首先回想矩阵乘法的如下性质:设 $AB=C$,则有

(1) C 的第 i 个列向量是把 A 的列向量组用 B 的第 i 列元素

为系数作线性组合得来的；

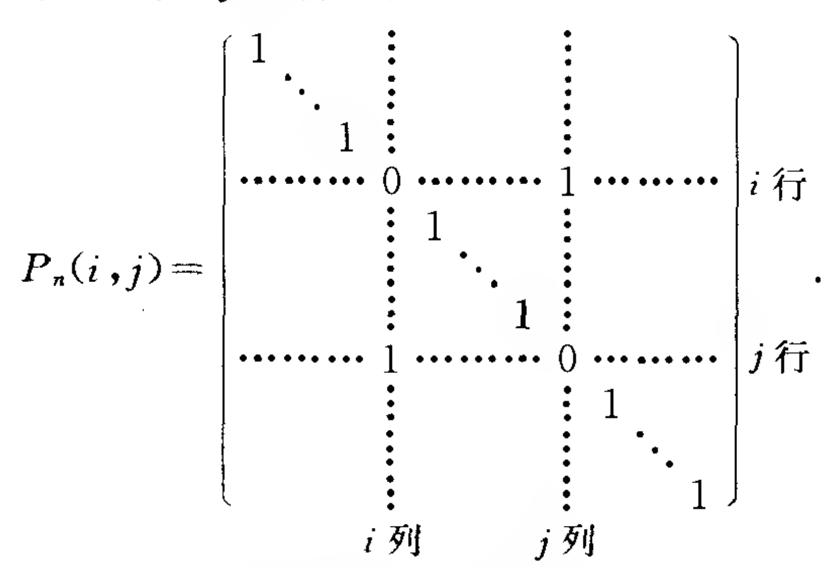
(2) C 的第 i 个行向量是把 B 的行向量组用 A 的第 i 行元素为系数作线性组合得来的.

性质(1)已在§1中指出了，而性质(2)可由性质(1)得到. 因为：C 的第 i 行为 C' 的第 i 列，而 $C'=(AB)'=B'A'$，故 C' 的第 i 列是 B' 的列向量组用 A' 的第 i 列元素作线性组合得出的. B' 的列向量组即 B 的行向量组，A' 的第 i 列即 A 的第 i 行，故知性质(2)正确.

定义 n 阶单位矩阵 E 经过一次初等行变换或初等列变换所得的矩阵称为 n 阶**初等矩阵**.

下面把初等矩阵分为三种类型分别写出来.

(1) 互换 E 的 i,j 两行，得



显然，互换 E 的 i,j 两列得到相同的结果.

(2) 把 E 的第 i 行乘以 $c\neq 0$，得

$$P_n(c\cdot i)=\begin{pmatrix} 1 & & & & & \\ & \ddots & & \vdots & & \\ & & 1 & \vdots & & \\ \cdots & \cdots & \cdots & c & \cdots & \cdots \\ & & & \vdots & 1 & \\ & & & \vdots & & \ddots \\ & & & \vdots & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ i\text{ 行} \\ \\ \\ \\ \end{matrix}.$$
$$\qquad\qquad i\text{ 列}$$

显然，把 E 的第 i 列乘以 c 得到相同的结果.

(3) 把 E 的第 j 行加上第 i 行的 k 倍,得

$$P_n(k\cdot i,j)=\begin{pmatrix} 1 & & \vdots & & \vdots & & \\ & \ddots & \vdots & & \vdots & & \\ \cdots\cdots & & 1 & \cdots & \vdots & \cdots\cdots & \\ & & \vdots & \ddots & \vdots & & \\ \cdots\cdots & & k & \cdots & 1 & \cdots\cdots & \\ & & \vdots & & \vdots & \ddots & \\ & & \vdots & & \vdots & & 1 \end{pmatrix}\begin{matrix} \\ \\ i\text{ 行} \\ \\ j\text{ 行} \\ \\ \\ \end{matrix}$$

i 列 j 列

(上面各矩阵的空白处均为零).

对 E 还可以作第四种初等变换:把第 j 列加上第 i 列的 k 倍,得

$$P_n'(k\cdot i,j)=\begin{pmatrix} 1 & & \vdots & & \vdots & & \\ & \ddots & \vdots & & \vdots & & \\ \cdots\cdots & & 1 & \cdots & k & \cdots\cdots & \\ & & \vdots & \ddots & \vdots & & \\ \cdots\cdots & & \vdots & \cdots & 1 & \cdots\cdots & \\ & & \vdots & & \vdots & \ddots & \\ & & \vdots & & \vdots & & 1 \end{pmatrix}\begin{matrix} \\ \\ i\text{ 行} \\ \\ j\text{ 行} \\ \\ \\ \end{matrix}.$$

i 列 j 列.

但是 $P_n'(k\cdot i,j)$ 可看作 $P_n(k\cdot j,i)$,即 E 的第 i 行加上第 j 行的 k 倍,所以它实际上属于类型(3).

命题 2.1 给定 $m\times n$ 矩阵 A,则有

(i) $P_m(i,j)A$ 为互换 A 的 i,j 两行;$AP_n(i,j)$ 为互换 A 的 i,j 两列.

(ii) $P_m(c\cdot i)A$ 为把 A 的第 i 行乘以 $c\neq 0$;$AP_n(c\cdot i)$ 为把 A 的第 i 列乘以 $c\neq 0$.

(iii) $P_m(k\cdot i,j)A$ 为把 A 的第 j 行加上第 i 行的 k 倍;$AP_n'(k\cdot i,j)$ 为把 A 的第 j 列加上第 i 列的 k 倍.

证 设 A 的列向量组为 $\alpha_1,\alpha_2,\cdots,\alpha_n$.

(i) $AP_n(i,j)$ 的第 k 个列向量为 A 的列向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 用 $P_n(i,j)$ 的第 k 列元素作线性组合所得到的. 当 $k\neq i,j$ 时即为 α_k;而当 $k=i$ 时为 α_j,当 $k=j$ 时为 α_i. 故相当于互换 A 的 i,j 两

列，其它不动. 同理可证 $P_m(i,j)A$ 为互换 A 的 i,j 两行，其它不动.

(ii) $AP_n(c\cdot i)$的第 k 个列向量是 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 用 $P_n(c\cdot i)$的第 k 列元素作线性组合所得. $k\neq i$ 时即为 α_k；而 $k=i$ 时为 $c\alpha_i$. 这相当于把 A 的第 i 列乘 c. $P_m(c\cdot i)A$ 的证法与此相同.

(iii) $AP'_n(k\cdot i,j)$的第 l 个列向量为 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 用 $P'_n(k\cdot i,j)$的第 l 列元素作线性组合所得. 当 $l\neq j$ 时即为 α_l；当 $l=j$ 时为 $k\alpha_i+\alpha_j$. 这相当于把 A 的第 j 列加上第 i 列的 k 倍. $P_m(k\cdot i,j)A$的证法相同. ▌

命题 2.1 把矩阵的初等行(列)变换归结为用某些初等矩阵左(右)乘该矩阵. 在第一章 § 3 中已指出，矩阵 A 可用初等行变换和列变换化为标准形 D. 由命题 2.1 可知，必存在初等矩阵 $P_1,P_2,\cdots,P_l$； $Q_1,Q_2,\cdots,Q_s$，使

$$P_1P_2\cdots P_lAQ_1Q_2\cdots Q_s=D,$$

其中 P_i 为 m 阶初等矩阵，Q_j 为 n 阶初等矩阵，而

$$D=\begin{pmatrix}1&&&&&\\&\ddots&&&0&\\&&1&&&\\&&&0&&\\&0&&&\ddots&\\&&&&0&\cdots\ 0\end{pmatrix},\quad\begin{pmatrix}1&&&&\\&\ddots&&&\\&&1&&\\&&&0&\\&&&&\ddots\\&&&&&0\end{pmatrix},\quad\begin{pmatrix}1&&&&&\\&\ddots&&&0&\\&&1&&&\\&&&0&&\\&&&&\ddots&\\&0&&&&0\\&&&&&\vdots\\&&&&&0\end{pmatrix}.$$

($n>m$ 时)　　　　($n=m$ 时)　　　　($n<m$ 时)

因为矩阵的初等变换是可逆的，矩阵 A 经若干次初等行、列变换化为 D，那么，矩阵 D 也可经若干次初等行、列变换化为 A. 因而，按命题 2.1，存在初等矩阵$\overline{P}_1,\overline{P}_2,\cdots,\overline{P}_l$； $\overline{Q}_1,\overline{Q}_2,\cdots,\overline{Q}_s$，使

$$\overline{P}_1\ \overline{P}_2\cdots\overline{P}_lD\,\overline{Q}_1\ \overline{Q}_2\cdots\overline{Q}_s=A.$$

现设 A 是一个 n 阶方阵. 如果 $r(A)=n$，则称 A 是一个**满秩**的 n 阶方阵. 一个满秩 n 阶方阵的标准形就是 n 阶单位矩阵 E. 故有

$$A=\overline{P}_1\ \overline{P}_2\cdots\overline{P}_l E\overline{Q}_1\ \overline{Q}_2\cdots\overline{Q}_s=\overline{P}_1\ \overline{P}_2\cdots\overline{P}_l\ \overline{Q}_1\ \overline{Q}_2\cdots\overline{Q}_s.$$

于是我们有如下的

命题 2.2 设 A 是 n 阶方阵，则 A 是满秩的当且仅当 A 可以表为 n 阶初等矩阵的乘积.

证 上面的讨论表明，如果 A 是满秩的，则 A 可以表为初等矩阵的乘积. 反之，如果 A 等于初等矩阵 $P_1,P_2,\cdots,P_s$ 的乘积，则

$$A=P_1P_2\cdots P_s=P_1P_2\cdots P_sE,$$

其中 E 为 n 阶单位矩阵. 这表明 A 可以由单位矩阵经过初等行变换得到. 而初等变换不改变矩阵的秩，所以 $r(A)=r(E)=n$，即 A 是满秩的. ▌

根据这一命题，所有 n 阶初等矩阵都是满秩的.

下面我们证明一个重要的事实.

定理 1 给定数域 K 上的矩阵 $A_{m\times n},B_{n\times s}$，那么，我们有

$$r(AB)\leqslant\min\{r(A),r(B)\}.$$

证 分两步证明.

(i) 先证 $r(AB)\leqslant r(B)$. 设 $AB=C$，则由本节开头对矩阵乘法所作的分析，C 的每个行向量都可被 B 的行向量组线性表示，从而都能被 B 的行向量组的极大线性无关部分组线性表示. 所以，C 的行向量组的极大线性无关部分组能被 B 的行向量组的极大线性无关部分组线性表示. 根据第一章命题 2.4 知，$r(C)\leqslant r(B)$.

(ii) 再证 $r(AB)\leqslant r(A)$.

$$r(AB)=r((AB)')=r(B'A')\leqslant r(A')=r(A)$$

(其中利用了(i)的结果).

综合(i),(ii)即得定理的证明. ▌

如果对一个 $m\times n$ 矩阵 A 左乘或右乘的是一个满秩方阵，那么，可以有更强的结果：

命题 2.3 设 A 是一个 $m\times n$ 矩阵，P 是 m 阶满秩方阵，Q 是 n 阶满秩方阵，则

$$r(PA)=r(AQ)=r(A).$$

即左乘或右乘满秩方阵后,矩阵的秩不变.

证 P 是满秩方阵,由命题 2.2,P 可表为初等矩阵的乘积:$P=P_1P_2\cdots P_k$. 故

$$PA=P_1P_2\cdots P_kA.$$

这说明 PA 可由 A 经 k 次初等行变换得到. A 的秩在初等行变换下不变,故

$$r(PA)=r(A).$$

同理可证 $r(AQ)=r(A)$. ▌

习 题 二

1. 做矩阵乘法:

(1) $$\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}\begin{pmatrix}-2&1&3&-1\\0&1&1&-1\\2&0&1&0\end{pmatrix};$$

$$\begin{pmatrix}-2&1&3&-1\\0&1&1&-1\\2&0&1&0\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{pmatrix}.$$

(2) $$\begin{pmatrix}1&0&0\\0&1&-2\\0&0&1\end{pmatrix}\begin{pmatrix}-2&1&3&-1\\0&1&1&-1\\2&0&1&0\end{pmatrix};$$

$$\begin{pmatrix}-2&1&3&-1\\0&1&1&-1\\2&0&1&0\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&-2&0&1\end{pmatrix}.$$

说明它们和矩阵初等变换的关系.

2. 求三阶方阵 P,Q,使

(1) $P\begin{pmatrix}1 & -1 & 2\\ 2 & 0 & 1\\ -1 & 1 & 0\end{pmatrix}=\begin{pmatrix}1 & -1 & 2\\ 0 & 2 & -3\\ 0 & 0 & 2\end{pmatrix}$;

(2) $\begin{pmatrix}1 & -1 & 2\\ 2 & 0 & 1\\ -1 & 1 & 0\end{pmatrix}Q=\begin{pmatrix}1 & 0 & 0\\ 2 & 2 & -3\\ -1 & 0 & 2\end{pmatrix}$.

3. 设 A 是一个 n 阶方阵，P 是一个 n 阶初等矩阵，$B=P'AP$. 试讨论 B 可由 A 经过什么样的初等变换而得到.

4. 设 A 是 $m\times n$ 矩阵，B 是 $n\times s$ 矩阵，且 $r(B)=r$. 证明：

(i) 存在初等矩阵 $Q_1,Q_2,\cdots,Q_k$，使

$$BQ_1Q_2\cdots Q_k=C=\begin{pmatrix}c_{11} & \cdots & c_{1r} & 0 & \cdots & 0\\ c_{21} & \cdots & c_{2r} & 0 & \cdots & 0\\ \vdots & & \vdots & \vdots & & \vdots\\ c_{n1} & \cdots & c_{nr} & 0 & \cdots & 0\end{pmatrix};$$

(ii) $r(AC)\leqslant r$;

(iii) 利用上述结果证明定理 1.

5. 把矩阵

$$\begin{pmatrix}a & 0\\ 0 & \frac{1}{a}\end{pmatrix}\quad(a\neq 0)$$

表成形式为

$$\begin{pmatrix}1 & x\\ 0 & 1\end{pmatrix}\quad\text{与}\quad\begin{pmatrix}1 & 0\\ x & 1\end{pmatrix}$$

的矩阵的乘积.

6. 设

$$A=\begin{pmatrix}a & b\\ c & d\end{pmatrix},\quad \begin{vmatrix}a & b\\ c & d\end{vmatrix}=1,$$

证明：A 可表成形式为

$$\begin{pmatrix}1 & x\\ 0 & 1\end{pmatrix}\quad\text{与}\quad\begin{pmatrix}1 & 0\\ x & 0\end{pmatrix}$$

的矩阵的乘积.

7. 给定 n 阶上三角矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & 0 & & a_{nn} \end{pmatrix},$$

若 $a_{ii}\neq 0(i=1,2,\cdots,n)$,证明:存在 n 阶方阵 B,使 $BA=E$.

8. 给定 n 阶方阵 A,证明:存在初等矩阵 $P_1,P_2,\cdots,P_k$,使 $P_1P_2\cdots P_kA$ 成为上三角矩阵.

9. 给定 n 阶方阵 A,若 $r(A)=n$,证明存在 n 阶方阵 B,使 $BA=E$.

10. 证明第一类初等矩阵 $P_n(i,j)$ 可以表示成第二、三类初等矩阵 $P_n(c\cdot i)$ 和 $P_n(k\cdot i,j)$ 的乘积.

§3 逆 矩 阵

在§1中我们介绍了矩阵的加法(同时就有了减法)和数乘、乘法运算. 应当指出,在矩阵中一般地是不能做除法运算的. 但在本节里我们将要指出,在一定条件下,可以有类似于数的除法的概念.

两个数 $a,b(b\neq 0)$ 相除可用 $a\cdot b^{-1}$ 或 $b^{-1}a$ 来表示. 能不能做除法,关键在于对每个 $b\neq 0$,是否能找到一个数 b^{-1},使 $b^{-1}b=bb^{-1}=1$. 在数的范围内这总是可以办到的. 但在矩阵的范围内就不一定了,只有在更强的条件下才能做到这一点.

定义 设 A 是一个 n 阶方阵,如果存在 n 阶方阵 B,使

$$BA=AB=E,$$

则称 B 是 A 的一个**逆矩阵**,此时 A 称为**可逆矩阵**.

一个矩阵 A 如果可逆,那么它的逆矩阵必定是唯一的. 因为:设 B,B_1 都是 A 的逆矩阵,我们有

$$B=BE=B(AB_1)=(BA)B_1=EB_1=B_1.$$

今后，我们将一个可逆矩阵 A 的逆矩阵记作 A^{-1}. 而对任意正整数 k，令

$$A^{-k}=(A^{-1})^k.$$

注意，只有对方阵才能谈逆矩阵的问题.

引进逆矩阵的概念有重要的意义. 例如，把线性方程组写成矩阵方程式

$$AX=B.$$

如果 A 是一个可逆方阵，则以 A^{-1}左乘（不能右乘，因为矩阵乘法不可交换，X 与 B 都是 $n\times 1$ 矩阵，所以在这里右乘甚至没有意义）等式两边，得

$$A^{-1}(AX)=(A^{-1}A)X=EX=X=A^{-1}B.$$

即

$$X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=A^{-1}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

这就把线性方程组的解求出来了. 这很像解一元一次方程 $ax=b$ 所采用的办法. 关于这个问题，我们将在第三章 § 2 再做详细和严格的讨论.

例 1 设

$$A=\begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 2 & -1 & 1 \end{pmatrix},\quad B=\begin{pmatrix} 2 & -1 & -1 \\ 3 & -1 & -2 \\ -1 & 1 & 1 \end{pmatrix}.$$

做矩阵乘法，不难验证

$$BA=AB=E.$$

故 A 可逆，且 $A^{-1}=B$.

例 2 在 § 2 中所定义的初等矩阵都是可逆矩阵. 具体地说，有

(i) $P_n(i,j)^{-1}=P_n(i,j)$;

(ii) $P_n(c\cdot i)^{-1}=P_n\left(\frac{1}{c}\cdot i\right) \quad (c\neq 0)$;

(iii) $P_n(k\cdot i,j)^{-1}=P_n(-k\cdot i,j)$.

上面的关系式既可以用矩阵乘法直接验证,也可以用命题 2.1 加以证明.例如证明(iii):因为

$$P_n(k\cdot i,j)\cdot P_n(-k\cdot i,j)=E.$$

等式右边表示对 n 阶单位矩阵 E 作两次初等行变换,第一次是把第 j 行加上第 i 行的 $-k$ 倍,第二次是把变换后的矩阵的第 j 行再加上第 i 行的 k 倍,结果还变成 E,故

$$P_n(k\cdot i,j)\cdot P_n(-k\cdot i,j)=E.$$

同理,有

$$P_n(-k\cdot i,j)P_n(k\cdot i,j)=E.$$

从而
$$P_n(k\cdot i,j)^{-1}=P_n(-k\cdot i,j).$$

(i)与(ii)请读者自行证明.

关于逆矩阵,有如下几个简单的事实.

命题 3.1 设 A,B 是两个可逆 n 阶方阵,则有

(i) $(A^{-1})^{-1}=A$;

(ii) AB 可逆,且 $(AB)^{-1}=B^{-1}A^{-1}$;

(iii) A' 可逆,且 $(A')^{-1}=(A^{-1})'$.

证 (i) 设 $A^{-1}=C$,则因 $CA=AC=E$,故 A 可看作 C 的逆矩阵,即 $C^{-1}=A$,亦即 $(A^{-1})^{-1}=A$.

(ii) 因为

$$(AB)(B^{-1}A^{-1})=A(BB^{-1})A^{-1}=AEA^{-1}=AA^{-1}=E,$$

$$(B^{-1}A^{-1})(AB)=B^{-1}(A^{-1}A)B=B^{-1}EB=B^{-1}B=E,$$

故 $(AB)^{-1}=B^{-1}A^{-1}$.

(iii) 因为

$$A'(A^{-1})'=(A^{-1}A)'=E'=E,$$

$$(A^{-1})'A'=(AA^{-1})'=E'=E,$$

故$(A')^{-1}=(A^{-1})'$. ▌

根据上述命题的(ii)可知:有限多个可逆方阵 $A_1,A_2,\cdots,A_s$ 的乘积还是可逆的,且

$$(A_1A_2\cdots A_s)^{-1}=A_s^{-1}\cdots A_2^{-1}A_1^{-1}.$$

特别地,如 A 是一个 n 阶可逆矩阵,则对任意正整数 k,A^k 也可逆,且

$$(A^k)^{-1}=(A^{-1})^k=A^{-k}.$$

定理 2 一个 n 阶方阵 A 可逆的充分必要条件是:A 是满秩的,即 $r(A)=n$.

证 (i) 必要性. 当 A 可逆时,存在 B,使

$$AB=E.$$

由定理 1,有

$$r(AB)\leqslant \min(r(A),r(B)).$$

因为 $r(AB)=r(E)=n$,故 $r(A)=n$.

(ii) 若 $r(A)=n$,则由命题 2.2,A 可表为初等矩阵的乘积. 初等矩阵是可逆的,它们的乘积也可逆,故 A 可逆. ▌

推论 对 n 阶方阵 A,如果存在一个 n 阶方阵 B,使下列等式之一成立:

$$AB=E \quad 或 \quad BA=E,$$

则 A 可逆,且 $A^{-1}=B$.

证 设 $AB=E$ 成立. 则由

$$n=r(E)=r(AB)\leqslant \min(r(A),r(B)),$$

知 A 满秩. 再根据定理 2 知,A 可逆,A^{-1}存在. 而

$$A^{-1}=A^{-1}E=A^{-1}(AB)=(A^{-1}A)B=EB=B.$$

当 $BA=E$ 成立时证法相同. ▌

这个推论说明:要验证 B 是否是 A 的逆矩阵,只要从左侧或右侧去乘看是否得到单位矩阵就可以了.

下面来介绍逆矩阵的计算方法. 设 A 是一个 n 阶可逆矩阵. 按定理 2, A 满秩. 从而,A 可以表示成初等矩阵的乘积

$$A=P_1P_2\cdots P_s.$$

以 $P_1^{-1},P_2^{-1},\cdots,P_s^{-1}$ 依次左乘等式两端，得

$$(P_s^{-1}\cdots P_2^{-1}P_1^{-1})A=E. \tag{1}$$

根据定理 2 的推论，有

$$A^{-1}=P_s^{-1}\cdots P_2^{-1}P_1^{-1}=P_s^{-1}\cdots P_2^{-1}P_1^{-1}E. \tag{2}$$

现在观察(1)式和(2)式. $P_1^{-1},\cdots,P_s^{-1}$ 都是初等矩阵，用它们依次左乘 A，相当于依次对 A 作初等行变换. (1)式表明这些初等行变换把 A 化成单位矩阵 E. 而(2)式表明：如果把这些初等行变换同时作用在单位矩阵 E 上，那么最后恰把 E 变成 A^{-1}. 由此得到逆矩阵的如下计算方法：

(i) 把 A 和 E 并排放在一起，排成一个 $n\times 2n$ 矩阵：(AE)；

(ii) 对上面的 $n\times 2n$ 矩阵做初等行变换(不能做列变换)，把左边的 A 化为 E 时，右边 E 的位置上所出来的就是 A^{-1}. 用图式表示，就是

$$(AE)\longrightarrow(EA^{-1}).$$

读者仿照上面的办法自己证明一下：如把 A 和 E 上下排列成一个 $2n\times n$ 矩阵

$$\begin{pmatrix} A \\ E \end{pmatrix},$$

对它做列初等变换(不能做行变换)，把上面的 A 化成 E，则此时下面 E 的位置上所出来的也是 A^{-1}

$$\begin{pmatrix} A \\ E \end{pmatrix}\longrightarrow\begin{pmatrix} E \\ A^{-1} \end{pmatrix}.$$

例 3 求

$$A=\begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 2 & -1 & 1 \end{pmatrix}$$

的逆矩阵.

解 给定 A，我们事先可能不知道它是否可逆. 但并不需要先

去判断它是否可逆，而可以直接利用上面介绍的办法进行计算. 如果 A 不可逆，那么做行初等变换，把 A 化成阶梯形后，其阶梯个数 $=r(A)<n$. 若 A 可逆，则 $r(A)=n$，此时先把 A 化为阶梯形，再把主对角线上元素全变为 1，再利用最后一行最后一列的 1 消去它顶上的其它数，又利用倒数第二行主对角线上的 1 消去它顶上的其它数，等等. 本例具体计算如下：

$$(AE)=\begin{pmatrix}1&0&1&1&0&0\\-1&1&1&0&1&0\\2&-1&1&0&0&1\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&1&1&0&0\\0&1&2&1&1&0\\0&-1&-1&-2&0&1\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1&0&1&1&0&0\\0&1&2&1&1&0\\0&0&1&-1&1&1\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&1&1&0&0\\0&1&0&3&-1&-2\\0&0&1&-1&1&1\end{pmatrix}$$

$$\longrightarrow\left(\begin{array}{ccc:ccc}1&0&0&2&-1&-1\\0&1&0&3&-1&-2\\0&0&1&-1&1&1\end{array}\right).$$

$$A^{-1}=\begin{pmatrix}2&-1&-1\\3&-1&-2\\-1&1&1\end{pmatrix}.$$

例 4 解矩阵方程

$$\begin{pmatrix}1&0&1\\-1&1&1\\2&-1&1\end{pmatrix}\begin{pmatrix}x_1&y_1\\x_2&y_2\\x_3&y_3\end{pmatrix}=\begin{pmatrix}1&1\\0&1\\-1&0\end{pmatrix}.$$

解 显然，有

$$\begin{pmatrix}x_1&y_1\\x_2&y_2\\x_3&y_3\end{pmatrix}=\begin{pmatrix}1&0&1\\-1&1&1\\2&-1&1\end{pmatrix}^{-1}\begin{pmatrix}1&1\\0&1\\-1&0\end{pmatrix}.$$

但处理这类问题实际上可以不必先计算逆矩阵. 设 A 是一个 n 阶可逆方阵，B 是 $n\times s$ 矩阵，要求一个 $n\times s$ 矩阵 X，使

$$AX=B$$

（显然，这样的 X 必存在，例如不难验证 $A^{-1}B$ 就是）. 已知有初等矩阵 $P_1,P_2,\cdots,P_l$，使

$$P_1P_2\cdots P_lA=E, \tag{3}$$

把这些初等矩阵左乘上面的矩阵方程，得

$$P_1P_2\cdots P_lAX=P_1P_2\cdots P_lB,$$

即

$$X=P_1P_2\cdots P_lB. \tag{4}$$

比较(3)，(4)两式可知：如果用一系列初等行变换把 A 化为 E，同时把这些初等行变换施加在 B 上，所得的结果即为 X 的解. 用图式表示，就是

$$(AB)\longrightarrow(EX).$$

例 4 具体计算如下

$$(AB)=\begin{pmatrix}1 & 0 & 1 & 1 & 1\\ -1 & 1 & 1 & 0 & 1\\ 2 & -1 & 1 & -1 & 0\end{pmatrix}\longrightarrow\begin{pmatrix}1 & 0 & 1 & 1 & 1\\ 0 & 1 & 2 & 1 & 2\\ 0 & -1 & -1 & -3 & -2\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1 & 0 & 1 & 1 & 1\\ 0 & 1 & 2 & 1 & 2\\ 0 & 0 & 1 & -2 & 0\end{pmatrix}\longrightarrow\left(\begin{array}{ccc:cc}1 & 0 & 0 & 3 & 1\\ 0 & 1 & 0 & 5 & 2\\ 0 & 0 & 1 & -2 & 0\end{array}\right).$$

故

$$\begin{pmatrix}x_1 & y_1\\ x_2 & y_2\\ x_3 & y_3\end{pmatrix}=\begin{pmatrix}3 & 1\\ 5 & 2\\ -2 & 0\end{pmatrix}.$$

当 X 和 B 都是 $n\times 1$ 矩阵时，$AX=B$ 就是 n 个未知量 n 个方程的线性方程组，而这时上面所述的办法就是第一章 §1 所讲的矩阵消元法. 读者可仿照上述办法导出解矩阵方程

$$XA=B$$

（其中 A 是 n 阶可逆方阵，X,B 是 $m\times n$ 矩阵）的方法.

习 题 三

1. 给定

$$A=\begin{pmatrix}1&3&2\\15&2&0\\4&2&1\end{pmatrix}.$$

(1) 证明：

$$A^{-1}=\begin{pmatrix}2&1&-4\\-15&7&30\\22&10&-43\end{pmatrix}.$$

(2) 利用上述结果解线性方程组

$$\begin{cases}x_1+3x_2+2x_3=b_1,\\15x_1+2x_2\qquad=b_2,\\4x_1+2x_2+\ x_3=b_3.\end{cases}$$

2. 证明：$P_n(i,j)^{-1}=P_n(i,j)$；$P_n(c\cdot i)^{-1}=P_n\left(\frac{1}{c}\cdot i\right)$.

3. 计算逆矩阵.

(1) $A=\begin{pmatrix}a&b\\c&d\end{pmatrix}$，$ad-bc=1$；

(2) $A=\begin{pmatrix}1&1&-1\\2&1&0\\1&-1&0\end{pmatrix}$；　(3) $A=\begin{pmatrix}2&2&3\\1&-1&0\\-1&2&1\end{pmatrix}$；

(4) $A=\begin{pmatrix}1&2&3&4\\2&3&1&2\\1&1&1&-1\\1&0&-2&-6\end{pmatrix}$；(5) $A=\begin{pmatrix}1&1&1&1\\1&1&-1&-1\\1&-1&1&-1\\1&-1&-1&1\end{pmatrix}$；

(6) $A=\begin{pmatrix}3&3&-4&-3\\0&6&1&1\\5&4&2&1\\2&3&3&2\end{pmatrix}$；(7) $A=\begin{pmatrix}1&3&-5&7\\0&1&2&-3\\0&0&1&2\\0&0&0&1\end{pmatrix}$；

(8) $A=\begin{pmatrix} 2 & 1 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ 5 & 7 & 1 & 8 \\ -1 & -3 & -1 & -6 \end{pmatrix}$；(9) $A=\begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & 3 & 1 & 4 \\ 2 & 7 & 6 & -1 \\ 1 & 2 & 2 & -1 \end{pmatrix}$；

(10) $A=\begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$.

4. 求方阵 X，使

(1) $\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} X=\begin{pmatrix} 4 & -6 \\ 2 & 1 \end{pmatrix}$；

(2) $\begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 2 \\ 1 & -1 & 0 \end{pmatrix} X=\begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix}$；

(3) $X\begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 2 \\ 1 & -1 & 0 \end{pmatrix}=\begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix}$；

(4) $\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix}_{n\times n} X=\begin{pmatrix} 2 & 1 & & & \\ 1 & 2 & 1 & & 0 \\ & 1 & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & 0 & & 1 & 2 \end{pmatrix}_{n\times n}$.

5. 设

$$A=\begin{pmatrix} 0 & a_1 & & & \\ \vdots & 0 & a_2 & & 0 \\ \vdots & \vdots & \ddots & \ddots & \\ 0 & \vdots & & \ddots & a_{n-1} \\ a_n & 0 & \cdots & \cdots & 0 \end{pmatrix},$$

其中 $a_i\neq 0(i=1,2,\cdots,n)$，求 A^{-1}.

6. 设 A 是 n 阶可逆方阵,证明:

(1) 若 A 对称(反对称),则 A^{-1}也对称(反对称);

(2) 若 A 是上(下)三角矩阵,则 A^{-1}也是上(下)三角矩阵.

7. 设 A 是一个 n 阶方阵,$A^k=0$. 证明:

$$(E-A)^{-1}=E+A+A^2+\cdots+A^{k-1}.$$

8. 设

$$f(\lambda)=a_0\lambda^m+a_1\lambda^{m-1}+\cdots+a_m \quad (a_m\neq 0).$$

若 A 是一个 n 阶方阵,且 $f(A)=0$,证明 A 可逆,且

$$A^{-1}=-\frac{1}{a_m}(a_0A^{m-1}+a_1A^{m-2}+\cdots+a_{m-1}E).$$

9. 设 B 为可逆 n 阶方阵,又设

$$U=\begin{pmatrix}u_1\\u_2\\\vdots\\u_n\end{pmatrix},\quad V=\begin{pmatrix}v_1\\v_2\\\vdots\\v_n\end{pmatrix}.$$

令 $A=B+UV'$. 证明:当 $\gamma=1+V'B^{-1}U\neq 0$ 时,

$$A^{-1}=B^{-1}-\frac{1}{\gamma}(B^{-1}U)(V'B^{-1}).$$

10. 证明一个上(下)三角矩阵主对角线上元素全不为零时必定可逆.

§4 矩阵的分块运算

考察两个 n 阶对角矩阵

$$A=\begin{pmatrix}a_1&&&\\&a_2&&\\&&\ddots&\\&&&a_n\end{pmatrix},\quad B=\begin{pmatrix}b_1&&&\\&b_2&&\\&&\ddots&\\&&&b_n\end{pmatrix}.$$

显然,它们的乘积是

$$AB=\begin{pmatrix} a_1b_1 & & & \\ & a_2b_2 & & \\ & & \ddots & \\ & & & a_nb_n \end{pmatrix}.$$

这说明，两个对角矩阵的乘积还是对角矩阵，而且恰好就是主对角线上的元素对应相乘. 因此，从矩阵乘法运算的角度来看，对角矩阵是最简单的一类矩阵.

下面来介绍一类稍复杂一点的矩阵. 先看一个例子. 考察

$$A=\begin{pmatrix} a_{11} & a_{12} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 & 0 \\ 0 & 0 & a_{33} & a_{34} & a_{35} \\ 0 & 0 & a_{43} & a_{44} & a_{45} \\ 0 & 0 & a_{53} & a_{54} & a_{55} \end{pmatrix},\quad B=\begin{pmatrix} b_{11} & b_{12} & 0 & 0 & 0 \\ b_{21} & b_{22} & 0 & 0 & 0 \\ 0 & 0 & b_{33} & b_{34} & b_{35} \\ 0 & 0 & b_{43} & b_{44} & b_{45} \\ 0 & 0 & b_{53} & b_{54} & b_{55} \end{pmatrix}.$$

如果令

$$A_1=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},\quad B_1=\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix},$$

$$A_2=\begin{pmatrix} a_{33} & a_{34} & a_{35} \\ a_{43} & a_{44} & a_{45} \\ a_{53} & a_{54} & a_{55} \end{pmatrix},\quad B_2=\begin{pmatrix} b_{33} & b_{34} & b_{35} \\ b_{43} & b_{44} & b_{45} \\ b_{53} & b_{54} & b_{55} \end{pmatrix}.$$

那么，上述两个五阶方阵可简写为

$$A=\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},\quad B=\begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix},$$

其中左下角的 0 表 3×2 零矩阵，右上角的 0 表 2×3 零矩阵. 这样表示之后，A，B 的样子很像对角矩阵，只是现在它们里面的“元素”不是数，而是小块矩阵. 不难验证：此时有

$$AB=\begin{pmatrix} A_1B_1 & 0 \\ 0 & A_2B_2 \end{pmatrix},$$

其中 A_1B_1 是两个矩阵的乘积，A_2B_2 也是两个矩阵的乘积，所以要注意它们乘积的次序，不能交换位置. 从这个例子可以看出，这种

矩阵在做乘法时也比较简单.

定义 称下列 n 阶方阵

$$A=\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{pmatrix}$$

为**准对角矩阵**,其中 $A_i(i=1,2,\cdots,s)$ 为 n_i 阶方阵,且 $n_1+n_2+\cdots+n_s=n$(除 A_i 的位置外,其它位置处全是小块零矩阵).

n 阶准对角矩阵有如下性质:

(1) 对两个同类型的 n 阶准对角矩阵

$$A=\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{pmatrix},\quad B=\begin{pmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_s \end{pmatrix}$$

(其中 $A_i,B_i(i=1,2,\cdots,s)$ 同为 n_i 阶方阵),有

$$AB=\begin{pmatrix} A_1B_1 & & & \\ & A_2B_2 & & \\ & & \ddots & \\ & & & A_sB_s \end{pmatrix};$$

(2) $r(A)=r(A_1)+r(A_2)+\cdots+r(A_s)$;

(3) A 可逆 $\Longleftrightarrow A_i(i=1,2,\cdots,s)$ 都可逆,且

$$A^{-1}=\begin{pmatrix} A_1^{-1} & & & \\ & A_2^{-1} & & \\ & & \ddots & \\ & & & A_s^{-1} \end{pmatrix}.$$

性质(1)是显然的,我们证明性质(2)及(3).

性质(2)的证明:在矩阵 A 中用初等变换分别把 $A_1,A_2,\cdots,A_s$ 化为标准形.在对 A_i 做变换时,对其它 A_j 没有影响.每个 A_i 位置所出来的标准形含有 $r(A_i)$ 个 1,再调换一下位置即得 A 的标准形,其中 1 的个数为

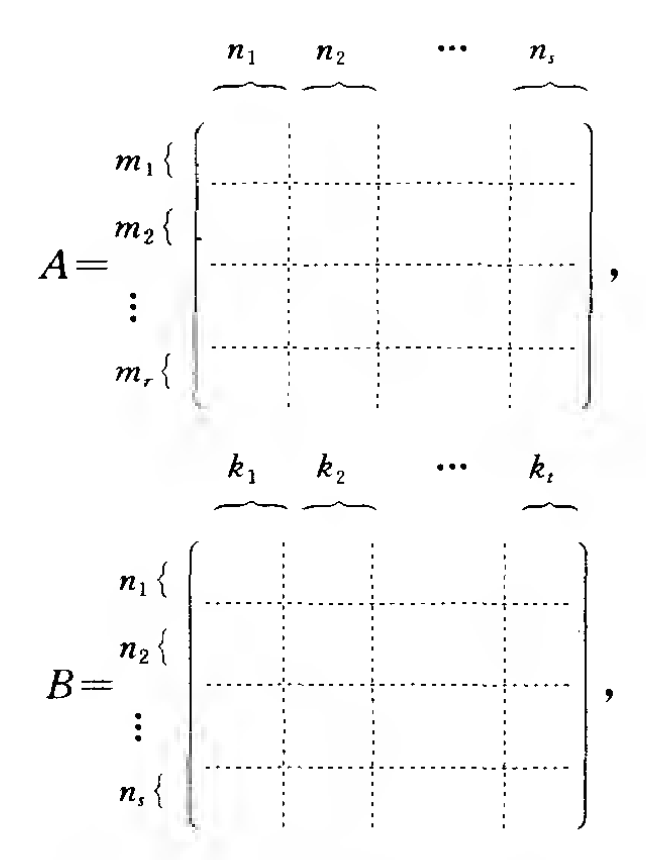
$$r(A)=r(A_1)+r(A_2)+\cdots+r(A_s).$$

性质(3)的证明:A 可逆 $\Longleftrightarrow r(A)=n$(参看定理 2).再由性质

(2)知，$r(A)=n$ 的充分必要条件是 $r(A_i)=n_i$，即 A_i 可逆. 而 A^{-1} 的上述表述式从性质(1)不难验证.

对于一般矩阵，我们可以对它进行"分块". 这样做在许多情况下能使问题简化.

设 A 是 $m\times n$ 矩阵，B 是 $n\times k$ 矩阵. 把它们按如下方式分割成小块：



即将 A 的行分割为 r 段，每段分别包含 $m_1,m_2,\cdots,m_r$ 个行，又将 A 的列分割为 s 段，每段分别包含 $n_1,n_2,\cdots,n_s$ 个列. 于是 A 可用小块矩阵表示如下

$$A=\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{pmatrix},$$

其中 A_{ij}为 $m_i\times n_j$ 矩阵. 对 B 做类似的分割，只是要求它的行的分割法和 A 的列的分割法相同(列的分割法没有限制). 于是 B 可表为

$$B=\begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1t} \\ B_{21} & B_{22} & \cdots & B_{2t} \\ \vdots & \vdots & & \vdots \\ B_{s1} & B_{s2} & \cdots & B_{st} \end{pmatrix},$$

其中 B_{ij}是 $n_i \times k_j$ 矩阵. 这种分割法称为**矩阵的分块**. 此时,设

$$AB=C,$$

则 C 有如下分块形式:

$$C=\begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1t} \\ C_{21} & C_{22} & \cdots & C_{2t} \\ \vdots & \vdots & & \vdots \\ C_{r1} & C_{r2} & \cdots & C_{rt} \end{pmatrix},$$

其中 C_{ij}是 $m_i \times k_j$ 矩阵,且

$$C_{ij} = \sum_{l=1}^{s} A_{il} B_{lj}.$$

它在形式上与用数作元素的普通矩阵的乘法法则一样,只是要注意现在 $A_{il}B_{lj}$是矩阵的乘法,一般不能交换次序.

如果用 E_m, E_n 分别代表 m, n 阶单位矩阵,那么,对任意 $m\times n$ 矩阵 X 和 $n\times m$ 矩阵 Y,下面分块矩阵

$$R=\begin{pmatrix} E_n & 0 \\ X & E_m \end{pmatrix}, \quad S=\begin{pmatrix} E_n & Y \\ 0 & E_m \end{pmatrix}$$

都是满秩的 $m+n$ 阶方阵(例如,对 R 做初等行变换,利用其前 n 行元素,即小块单位矩阵 E_n 中元素,可把 X 处元素化为 0,于是 R 经初等行变换化为 $m+n$ 阶单位矩阵 E_{m+n};同理,S 可经初等列变换化为 E_{m+n}).

现设 A 为一 n 阶满秩方阵,B 为 $n\times s$ 矩阵,C 为 $m\times n$ 矩阵,D 为 $m\times s$ 矩阵. 考察如下分块矩阵

$$T=\begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

现在取

$$R=\begin{pmatrix} E_n & 0 \\ -CA^{-1} & E_m \end{pmatrix},$$

那么，我们有

$$RT=\begin{pmatrix} A & B \\ 0 & D-CA^{-1}B \end{pmatrix}.$$

因为 R 为满秩 $m+n$ 阶方阵，根据命题 2.3，我们有

$$r(T)=r(RT).$$

因为 A 满秩，矩阵(AB)可经初等行变换和列变换化为$(E_n\ 0)$，由此知

$$r(RT)=r(A)+r(D-CA^{-1}B)=n+r(D-CA^{-1}B).$$

因此，$r(T)=r(A)=n$ 的充分必要条件是

$$r(D-CA^{-1}B)=0,$$

即
$$D=CA^{-1}B.$$

同样地，如果取

$$S=\begin{pmatrix} E_n & -A^{-1}B \\ 0 & E_m \end{pmatrix},$$

那么，我们有

$$TS=\begin{pmatrix} A & 0 \\ C & D-CA^{-1}B \end{pmatrix},$$

这里同样有 $r(T)=r(TS)$.

上面利用 R(或 S)左(或右)乘分块矩阵 T，有点像对一个二阶方阵作初等行(或列)变换，把其右下角(或左上角)元素化为 0. 这种方法显然可以类似地推广到一般分块矩阵上去. 本节习题 4 及习题 7 说明这种方法对讨论方阵的逆矩阵是有用的. 实际上，这种方法还有其它许多应用，此处不再作进一步的讨论了.

习 题 四

1. 将下列各题中的矩阵分块后再做运算：

(1) $\begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 1 & 0 & 3 & -1 & 1 \\ -1 & 1 & 2 & 0 & 0 \\ 2 & 1 & 1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 3 & 1 & 1 & -1 & 1 \\ -2 & 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 & 1 \end{pmatrix}$;

(2) $\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 & 1 & 0 \\ 0 & 0 & 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & 0 & 0 & -3 \end{pmatrix}^5$;

(3) $\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^5$.

2. 求下面矩阵的逆矩阵

$$A=\begin{pmatrix} 1 & 1 & -1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 & 3 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & 1 \end{pmatrix}.$$

3. 设

$$X=\begin{pmatrix} 0 & A \\ C & 0 \end{pmatrix},$$

其中 A,C 是两个可逆方阵.设已知 A^{-1},C^{-1},求 X^{-1}.

4. 设

$$D=\begin{pmatrix} A & 0 \\ C & B \end{pmatrix},$$

其中 A,B 是两个可逆方阵.设已知 A^{-1},B^{-1},求 D^{-1}.

5. 设 A,B 分别为 m,n 阶方阵.如果存在 m,n 阶可逆方阵 T_1,T_2,使 $T_1^{-1}AT_1$ 和 $T_2^{-1}BT_2$ 均为对角矩阵,试证:存在 $m+n$ 阶可逆方阵 T,使

$$T^{-1}\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}T$$

为对角矩阵.

6. 将一个 $m\times n$ 矩阵 R 分块

$$R=\begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

其中 A 为 k 阶可逆方阵.设

$$S=\begin{pmatrix} A & B \\ 0 & D-CA^{-1}B \end{pmatrix},$$

证明:$r(R)=r(S)$.

7. 设将一个 n 阶可逆方阵 R 分块

$$R=\begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

其中 A 为 k 阶可逆方阵.证明:

(1) $D_1=D-CA^{-1}B$ 为 $n-k$ 阶可逆方阵;

(2) 设已知 A^{-1},D_1^{-1},求 R^{-1}.

(提示:找一个分块矩阵 Q,使

$$RQ=\begin{pmatrix} A & 0 \\ C & D_1 \end{pmatrix},$$

再利用题 4 的结果.)

第三章　行　列　式

本章介绍 n 阶行列式的定义和基本性质. 行列式不但是研究线性方程组和矩阵的有用工具,而且在许多理论和实际应用问题中,它也发挥着重要的作用. 因此,行列式是线性代数中一个必不可少的基本概念.

§1　n 阶行列式的定义及性质

在日常生活中,人们用身高、体重等数据来描述一个人在形体方面的特征. 在物理学中,用物体的体积、质量等数据来刻画该物体的物理属性,这种方法也被用来研究矩阵. 矩阵是一个长方表格,它本身不是数,但我们可以用某些数据来刻画它的某种特征. 本章阐述的行列式概念,就是用来刻画数域 K 上一个 n 阶方阵的某种特征的一个重要数据.

行列式的定义

考察数域 K 上全体 n 阶方阵所成的集合 $M_n(K)$. 从集合 $M_n(K)$ 到数域 K 的一个映射 f 称为定义在 $M_n(K)$ 上的一个函数. 因此,$M_n(K)$ 上一个函数就是一个给定的法则,依照这一法则,K 上每个 n 阶方阵 A 对应于 K 内一个唯一确定的数 $f(A)$.

例如,设 $A=(a_{ij})\in M_n(K)$,我们定义 $f(A)=a_{11}$,即每个 n 阶方阵 A 在法则 f 下对应于其第一行第一列元素 a_{11},f 就是 $M_n(K)$ 上的一个函数. 由此看来,$M_n(K)$ 上的函数是很多的. 显然,并不是 $M_n(K)$ 上随便一个函数都有研究价值. 下面我们介绍 $M_n(K)$ 上具有某种特定属性的函数,它将成为研究 n 阶方阵的重

要工具.

为了使下面的阐述较为简明、清楚,我们在本章中将使用一些特定的记号.设 A 是数域 K 上一个 n 阶方阵,其行向量组为 α_1,$\alpha_2,\cdots,\alpha_n$(写成横排形式),列向量组为 $\beta_1,\beta_2,\cdots,\beta_n$(写成竖列形式),我们根据行文的需要把 A 写成

$$A=\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix}\quad 或\quad A=(\beta_1,\beta_2,\cdots,\beta_n).$$

如果我们只研究 A 的第 i 行或第 j 列,就写

$$A=\begin{pmatrix}\vdots\\ \alpha_i\\ \vdots\end{pmatrix}\quad 或\quad A=(\cdots,\beta_j,\cdots),$$

把不讨论的行(列)用省略号代替.于是 $M_n(K)$上一个函数 $f(A)$ 可以写成

$$f(A)=f\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix}=f\begin{pmatrix}\vdots\\ \alpha_i\\ \vdots\end{pmatrix}$$

或

$$f(A)=f(\beta_1,\beta_2,\cdots,\beta_n)=f(\cdots,\beta_j,\cdots).$$

定义 设 f 是定义在 $M_n(K)$上的一个函数,满足如下条件:对 K^n 中任意向量 $\alpha_1,\alpha_2,\cdots,\alpha_n,\alpha$(写成横排形式)以及 K 中任意数 k,都有

$$f\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_i+\alpha\\ \vdots\\ \alpha_n\end{pmatrix}=f\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_i\\ \vdots\\ \alpha_n\end{pmatrix}+f\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha\\ \vdots\\ \alpha_n\end{pmatrix},\quad f\begin{pmatrix}\alpha_1\\ \vdots\\ k\alpha_i\\ \vdots\\ \alpha_n\end{pmatrix}=kf\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_i\\ \vdots\\ \alpha_n\end{pmatrix}$$

（这里 $i=1,2,\cdots,n$），则 f 称为 $M_n(K)$ 上一个**行线性函数**.

设 g 是定义在 $M_n(K)$ 上一个函数，满足如下条件：对 K^n 中任意向量 $\beta_1,\beta_2,\cdots,\beta_n,\beta$（写成竖列形式）以及 K 中任意数 k，都有

$$\begin{aligned}&g(\beta_1,\cdots,\beta_j+\beta,\cdots,\beta_n)\\&\quad=g(\beta_1,\cdots,\beta_j,\cdots,\beta_n)+g(\beta_1,\cdots,\beta,\cdots,\beta_n),\\&\quad g(\beta_1,\cdots,k\beta_j,\cdots,\beta_n)=kg(\beta_1,\cdots,\beta_j,\cdots,\beta_n)\end{aligned}$$

（这里 $j=1,2,\cdots,n$），则称 g 为 $M_n(K)$ 上一个**列线性函数**.

例如，考察 $M_2(K)$. 设

$$A=\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}\in M_2(K).$$

定义 $f(A)=a_{11}a_{22}-a_{12}a_{21}$. 容易验证，$f$ 是 $M_2(K)$ 上一个行线性函数，也是一个列线性函数（请读者作为习题，按上面的定义自行验证）.

如果 $M_n(K)$ 上一个列线性函数 f 满足如下条件：当 $A\in M_n(K)$ 有两列元素相同时，必有 $f(A)=0$，则 f 称为**反对称**的列线性函数.

当然，我们可以类似地定义反对称的行线性函数，这就不再重复说明了. 读者容易看出，上面定义的 $M_2(K)$ 内的列（行）线性函数是反对称的.

命题 1.1 设 f 是 $M_n(K)$ 上的反对称列线性函数，那么，下面命题成立：

(i) 设将 $A\in M_n(K)$ 的 i,j 两列互换得出方阵 B，则 $f(B)=-f(A)$；

(ii) 设将 $A\in M_n(K)$ 的第 j 列加上其第 i 列的 k 倍（k 为 K 中任意取定的数）得出方阵 B，则有 $f(B)=f(A)$.

证 设 $A=(\cdots,\beta_i,\cdots,\beta_j,\cdots)$. 由于 f 是反对称的列线性函数，我们有

$$\begin{aligned}0&=f(\cdots,\beta_i+\beta_j,\cdots,\beta_i+\beta_j,\cdots)\\&=f(\cdots,\beta_i,\cdots,\beta_i,\cdots)+f(\cdots,\beta_i,\cdots,\beta_j,\cdots)\end{aligned}$$

$$+ f(\cdots,\beta_j,\cdots,\beta_i,\cdots) + f(\cdots,\beta_j,\cdots,\beta_j,\cdots)$$
$$= f(\cdots,\beta_i,\cdots,\beta_j,\cdots) + f(\cdots,\beta_j,\cdots,\beta_i,\cdots).$$

于是

$$f(A) = f(\cdots,\beta_i,\cdots,\beta_j,\cdots) = - f(\cdots,\beta_j,\cdots,\beta_i,\cdots)$$
$$= - f(B).$$

同样地,我们有

$$f(B) = f(\cdots,\beta_i,\cdots,\beta_j + k\beta_i,\cdots)$$
$$= f(\cdots,\beta_i,\cdots,\beta_j,\cdots) + f(\cdots,\beta_i,\cdots,k\beta_i,\cdots)$$
$$= f(A) + kf(\cdots,\beta_i,\cdots,\beta_i,\cdots) = f(A).$$ ▌

推论 设 f,g 是 $M_n(K)$上两个反对称列线性函数,且对某个 $A\in M_n(K)$有 $f(A)=g(A)$. 设 A 经有限次初等列变换变为方阵 B,则仍有 $f(B)=g(B)$.

证 显然只需考虑 B 是 A 做一次初等列变换得出的方阵就可以了. 下面分别讨论三种初等列变换.

(i) 设互换 A 的 i,j 两列变为 B,则按命题 1.1,有

$$f(B) = - f(A),\quad g(B) = - g(A),$$

从而 $f(B)=g(B)$.

(ii) 设将 A 的第 j 列乘以 K 内非零数 k 得出方阵 B,则

$$f(B) = f(\cdots,k\beta_j,\cdots) = kf(\cdots,\beta_j,\cdots) = kf(A).$$

同理,$g(B)=kg(A)$,于是 $f(B)=g(B)$.

(iii) 设将 A 的第 j 列加上第 i 列的 k 倍得出方阵 B,则按命题 1.1,有

$$f(B) = f(A) = g(A) = g(B).$$ ▌

下面给出本节的基本概念.

定义 设 f 是 $M_n(K)$上一个列线性函数且满足如下条件:

(i) 如果 $A\in M_n(K)$不满秩,则 $f(A)=0$;

(ii) 对 $M_n(K)$内单位矩阵 E,有 $f(E)=1$.

则 f 称为 $M_n(K)$上一个**行列式函数**.

例如,在 $M_2(K)$上定义函数 f 如下:若

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

则 $f(A)=a_{11}a_{22}-a_{12}a_{21}$. 那么，容易验证 f 是 $M_2(K)$上的一个行列式函数(具体验证留给读者作为练习).

容易看出，如果 f 是 $M_n(K)$上的行列式函数，那么当 A 有两列元素相等时，$r(A)<n$，故 $f(A)=0$. 因而，行列式函数必是反对称列线性函数.

命题 1.2 $M_n(K)$上的行列式函数是唯一的.

证 设 f 与 g 是 $M_n(K)$上两个行列式函数，我们需要证明：对任意 $A\in M_n(K)$，有 $f(A)=g(A)$.

(i) 如果 $r(A)<n$，那么按定义有 $f(A)=g(A)=0$.

(ii) 如果 $r(A)=n$，按第二章命题 2.2，A 可表为 n 阶初等矩阵 $P_1,\cdots,P_m$ 的乘积：

$$A = P_1\cdots P_m = EP_1\cdots P_m.$$

上式表明 A 可由单位矩阵 E 做 m 次初等列变换得出. 因为 f,g 均为 $M_n(K)$上反对称列线性函数，按命题 1.1 的推论，由 $f(E)=g(E)=1$ 可推出

$$f(A) = g(A). \blacksquare$$

上面我们给出了行列式函数的定义，并证明这种函数是唯一的. 但 $M_n(K)$上是否确实存在这样的函数，还是一个未解决的问题. 现在我们借助数学归纳法，具体地给出 $M_n(K)$上一个函数 $\det(A)$(det 是英文"行列式"(determinant)的前三个字母)，并证明它是一个行列式函数.

给定 $A\in M_n(K)$，我们用 $A\binom{i}{j}$表示划去 A 的第 i 行和第 j 列后所剩的 $n-1$ 阶方阵. 因此，$A\binom{i}{j}\in M_{n-1}(K)$.

定义 设 $A=(a_{11})\in M_1(K)$，定义

$$\det(A) = a_{11}.$$

设在集合 $M_{n-1}(K)$内函数 $\det(A)$已经定义. 那么，对

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \in M_n(K),$$

定义

$$\det(A) = a_{11}\det A\binom{1}{1} - a_{12}\det A\binom{1}{2} + \cdots + (-1)^{n+1}a_{1n}\det A\binom{1}{n}$$

$$= \sum_{k=1}^{n}(-1)^{k+1}a_{1k}\det A\binom{1}{k}.$$

这样，对任意正整数 n，在 $M_n(K)$ 内函数 $\det(A)$ 都已有定义. 下面为了书写时简单一些，我们用记号 $|A|$ 来代表 $\det(A)$. 如果 $A=(a_{ij})\in M_n(K)$，那么我们可以写成

$$\det(A) = |A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

现在具体地给出 $n=2,3$ 时 $\det(A)=|A|$ 的表达式.

$n=2$ 时，令

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

现在 $A\binom{1}{1}=(a_{22})$，$A\binom{1}{2}=(a_{21})$，故 $\det A\binom{1}{1}=a_{22}$，$\det A\binom{1}{2}=a_{21}$. 于是，按定义

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

它恰好是把二阶方阵 A（正方形表格）两条对角线上的元素相乘再相减.

$n=3$ 时，令

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

现在

$$A\binom{1}{1}=\begin{pmatrix}a_{22} & a_{23}\\ a_{32} & a_{32}\end{pmatrix},\quad A\binom{1}{2}=\begin{pmatrix}a_{21} & a_{23}\\ a_{31} & a_{33}\end{pmatrix},\quad A\binom{1}{3}=\begin{pmatrix}a_{21} & a_{22}\\ a_{31} & a_{32}\end{pmatrix}.$$

相应的函数值是

$$\left|A\binom{1}{1}\right|=a_{22}a_{33}-a_{23}a_{32};$$

$$\left|A\binom{1}{2}\right|=a_{21}a_{33}-a_{23}a_{31};$$

$$\left|A\binom{1}{3}\right|=a_{21}a_{32}-a_{22}a_{31}.$$

于是,按定义

$$\begin{aligned}\det(A)&=\begin{vmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}\\ &=a_{11}\left|A\binom{1}{1}\right|-a_{12}\left|A\binom{1}{2}\right|+a_{13}\left|A\binom{1}{3}\right|\\ &=a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}\\ &\quad-a_{13}a_{22}a_{31}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}.\end{aligned}$$

所以,对一个三阶方阵 A,函数 $\det(A)$的表达式中出现 6 项.

对一般 n 阶方阵 A,函数 $\det(A)$的表达式显然比较长,我们将在本章最后一节来阐述这个表达式.现在先来讨论一个简单情况.

命题 1.3 我们有

$$\begin{vmatrix}a_{11} & & & 0\\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix}=a_{11}a_{22}\cdots a_{nn}.$$

证 对 n 作数学归纳法. $n=1$ 时显然成立.设上面公式在 $n-1$ 阶方阵时已成立,那么对 n 阶方阵

$$A=\begin{pmatrix}a_{11} & & & 0\\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix},$$

因为按归纳假设有

$$\det A\binom{1}{1} = \begin{vmatrix} a_{22} & & 0 \\ \vdots & \ddots & \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{22}\cdots a_{nn},$$

于是按定义有

$$\det(A) = a_{11}\det A\binom{1}{1} = a_{11}a_{22}\cdots a_{nn}. \blacksquare$$

推论 $\det(E)=1$.

证 这是命题 1.3 的直接推论. ▌

这表明函数 $\det(A)$满足行列式函数定义中的最后一个条件. 下面来证明它也满足其它条件.

1. 首先我们证明 $\det(A)$是 $M_n(K)$内的列线性函数. 设 $A\in M_n(K)$，且 A 的第 j 列为两个向量之和 $\lambda\alpha+\mu\beta(\lambda,\mu\in K)$，即设 $A=(\cdots,\lambda\alpha+\mu\beta,\cdots)$，令

$$A_1 = (\cdots,\alpha,\cdots),\quad A_2 = (\cdots,\beta,\cdots),$$

上面用省略号代表的列向量均保持不动. 我们来证明：

$$|A| = \lambda|A_1| + \mu|A_2|.$$

设 $A=(a_{ik})$，其第 j 个列向量是

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} = \lambda\alpha + \mu\beta = \lambda\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \mu\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} \lambda a_1 + \mu b_1 \\ \lambda a_2 + \mu b_2 \\ \vdots \\ \lambda a_n + \mu b_n \end{pmatrix}.$$

现在对 n 作数学归纳法. $n=1$ 时，

$$|A| = \lambda a_1 + \mu b_1 = \lambda|A_1| + \mu|A_2|,$$

结论成立. 设 A 为 $n-1$ 阶方阵时我们需要的公式已成立，则当 A 为 n 阶方阵时，

$$|A| = \sum_{k\neq j}(-1)^{k+1}a_{1k}\left|A\binom{1}{k}\right| + (-1)^{j+1}(\lambda a_1 + \mu b_1)\left|A\binom{1}{j}\right|.$$

当 $k\neq j$ 时，$A\binom{1}{k}$有一列向量为 $\lambda\alpha'+\mu\beta'$（去掉 α,β 的第一分量 a_1，b_1 得出 α',β'），按归纳假设，有

$$\left|A\binom{1}{k}\right| = \lambda\left|A_1\binom{1}{k}\right| + \mu\left|A_2\binom{1}{k}\right|.$$

而$|A\binom{1}{j}|=|A_1\binom{1}{j}|=|A_2\binom{1}{j}|$，故

$$\begin{aligned}|A|=&\sum_{k\neq j}(-1)^{k+1}a_{1k}(\lambda|A_1\binom{1}{k}|+\mu|A_2\binom{1}{k}|)\\&+(-1)^{j+1}\lambda a_1|A_1\binom{1}{j}|+(-1)^{j+1}\mu b_1|A_2\binom{1}{j}|\\=&\lambda\Big(\sum_{k\neq j}(-1)^{k+1}a_{1k}|A_1\binom{1}{k}|+(-1)^{j+1}a_1|A_1\binom{1}{j}|\Big)\\&+\mu\Big(\sum_{k\neq j}(-1)^{k+1}a_{1k}|A_2\binom{1}{k}|+(-1)^{j+1}b_1|A_2\binom{1}{j}|\Big)\\=&\lambda|A_1|+\mu|A_2|.\end{aligned}$$

在上面的公式中，如分别令$\lambda=\mu=1$和$\mu=0$代入，即知$\det(A)$为$M_n(K)$内的列线性函数.

现在我们来证明：

2. $\det(A)$是反对称列线性函数.

对n作数学归纳法. $n=2$时，A有两列元素相同，只能是

$$A=\begin{pmatrix}a_1 & a_1\\ a_2 & a_2\end{pmatrix},$$

因此，$\det(A)=a_1a_2-a_1a_2=0$，即A为反对称列线性函数. 设对有两列相同的$n-1$阶方阵A，$\det(A)=0$. 我们来证明：当A是i,j两列相同的n阶方阵时，也有$\det(A)=0$. 设$A=(a_{st})$. $k\neq i,j$时，$A\binom{1}{k}$是有两列元素相等的$n-1$阶方阵，按归纳假设，$|A\binom{1}{k}|=0$. 因而现在有

$$|A|=(-1)^{i+1}a_{1i}|A\binom{1}{i}|+(-1)^{j+1}a_{1j}|A\binom{1}{j}|.$$

(i) 若$j=i\pm1$，即i,j为相邻两列，那么有$|A\binom{1}{i}|=|A\binom{1}{j}|$及$a_{1i}=a_{1j}$，此时显见$|A|=0$.

(ii) 若$j\neq i\pm1$，不妨设$j=i+s(s\geqslant2)$. 根据归纳假设，$\det(A)$为$M_{n-1}(K)$内反对称列线性函数，根据命题1.1，互换$|A\binom{1}{j}|$两列的位置时，其值变号. 而将$A\binom{1}{j}$的第i列与其右边的列经$s-1$次相邻两列的互换之后变为$A\binom{1}{i}$，故

$$|A\binom{1}{j}|=(-1)^{s-1}|A\binom{1}{i}|,$$

代入上面$|A|$的表达式，得(注意 $a_{1j}=a_{1i}$)

$$
\begin{aligned}
|A| &= (-1)^{i+1}a_{1i}\left|A\binom{1}{i}\right| + (-1)^{j+1}a_{1i}(-1)^{s-1}\left|A\binom{1}{i}\right| \\
&= \left((-1)^{i+1} + (-1)^{i+2s}\right)a_{1i}\left|A\binom{1}{i}\right| = 0.
\end{aligned}
$$

这样，我们已证明：对有两列元素相同的 n 阶方阵 A，$\det(A)=0$，即 $\det(A)$ 为 $M_n(K)$ 内反对称列线性函数.

3. 设 $A\in M_n(K)$，且 A 有一列元素全为 0，则必有 $\det(A)=0$.

对 n 作数学归纳法. $n=1$ 时，有 $A=(0)$，故 $\det(A)=0$. 设对有一列元素全为 0 的 $n-1$ 阶方阵 A，有 $\det(A)=0$，那么，当 A 是第 j 列元素全为 0 的 n 阶方阵时，我们有 $\left|A\binom{1}{k}\right|=0(k\neq j)$. 这是因为 $A\binom{1}{k}$是有一列元素全为 0 的 $n-1$ 阶方阵. 又注意 A 的第一行 j 列元素 $a_{1j}=0$，故

$$|A| = (-1)^{j+1}a_{1j}\left|A\binom{1}{j}\right| = 0.$$

命题 1.4 $\det(A)$是 $M_n(K)$内唯一的行列式函数.

证 我们已知 $\det(A)$是 $M_n(K)$上的列线性函数，且 $\det(E)=1$. 下面只需证明：当 $r(A)<n$ 时，$\det(A)=0$. 如果 A 不满秩，可设 A 的第 j 列为其它各列的线性组合. 把其余各列乘以适当倍数加到第 j 列，可使第 j 列变为 0，设所得方阵为 B. 因为 $\det(A)$为反对称列线性函数，根据命题 1.1，有 $\det(A)=\det(B)$. 但 B 有一列元素全为 0，所以 $\det(B)=0$，即得

$$\det(A) = 0. \blacksquare$$

给定 $A=(a_{ij})\in M_n(K)$，我们称数

$$\det(A) = |A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

为方阵 A 的**行列式**，有时简称为一个 n **阶行列式**. 这个数是刻画方阵 A 的某种特征的重要工具，在数学以至自然科学和工程技术各领域都有广泛的应用.

现在对 det(A)的性质作进一步的探讨.

命题 1.5 det(A)是 $M_n(K)$上的行线性函数.

证 设 $A=(a_{ij})$, A 的第 k 行为两个向量 $\lambda\alpha$ 与 $\mu\beta$ 之和,即设

$$\begin{aligned}(a_{k1},\cdots,a_{kn}) &= \lambda\alpha+\mu\beta=\lambda(a_1,\cdots,a_n)+\mu(b_1,\cdots,b_n)\\ &=(\lambda a_1+\mu b_1,\cdots,\lambda a_n+\mu b_n).\end{aligned}$$

又令 A_1, A_2 表示将 A 的第 k 个行向量分别换成 α,β 后所得的 n 阶方阵. 我们只要证明:对任意 $\lambda,\mu\in K$,有

$$|A|=\lambda|A_1|+\mu|A_2|$$

就可以了.

对 n 作数学归纳法. $n=1$ 时, $A=(\lambda a_1+\mu b_1)$,故

$$|A|=\lambda a_1+\mu b_1=\lambda|A_1|+\mu|A_2|.$$

设命题对 $n-1$ 阶方阵成立,我们来证明它对 n 阶方阵也成立.

若 $k=1$,那么按 det(A)的定义,我们有

$$\begin{aligned}|A| &= \sum_{i=1}^{n}(-1)^{i+1}a_{1i}\left|A\binom{1}{i}\right| = \sum_{i=1}^{n}(-1)^{i+1}(\lambda a_i+\mu b_i)\left|A\binom{1}{i}\right|\\ &= \lambda\sum_{i=1}^{n}(-1)^{i+1}a_i\left|A\binom{1}{i}\right|+\mu\sum_{i=1}^{n}(-1)^{i+1}b_i\left|A\binom{1}{i}\right|\\ &= \lambda|A_1|+\mu|A_2|,\end{aligned}$$

其中用到了 $A\binom{1}{i}=A_1\binom{1}{i}=A_2\binom{1}{i}$这一事实.

若 $k>1$,现在按归纳假设,有 $\left|A\binom{1}{i}\right|=\lambda\left|A_1\binom{1}{i}\right|+\mu\left|A\binom{1}{i}\right|$,于是

$$\begin{aligned}|A| &= \sum_{i=1}^{n}(-1)^{i+1}a_{1i}\left|A\binom{1}{i}\right|\\ &= \lambda\sum_{i=1}^{n}(-1)^{i+1}a_{1i}\left|A_1\binom{1}{i}\right|+\mu\sum_{i=1}^{n}(-1)^{i+1}a_{1i}\left|A_2\binom{1}{i}\right|\\ &= \lambda|A_1|+\mu|A_2|.\end{aligned}$$

现在分别以 $\lambda=\mu=1$ 及 $\mu=0$ 代入,即知 det(A)为 $M_n(K)$上的行线性函数. ▌

最后,我们来证明:方阵 A 和它的转置矩阵 A'的行列式相同.

命题 1.6 $\det(A')=\det(A)$.

证 定义 $M_n(K)$ 上的函数 f 如下:对任意 K 上 n 阶方阵 A, $f(A)=\det(A')$. 按命题 1.5,$\det(A)$ 为行线性函数,从而 $f(A)$ 为列线性函数(因为 A 的列为 A' 的行),且当 $r(A)<n$ 时,$r(A')=r(A)<n$,故 $f(A)=\det(A')=0$. 显然又有 $f(E)=\det(E)=1$,于是 $f(A)$ 是 $M_n(K)$ 上的行列式函数. 根据命题 1.4,$\det(A)$ 是 $M_n(K)$ 上唯一的行列式函数,故 $f(A)=\det(A)$,亦即

$$\det(A') = \det(A). \blacksquare$$

命题 1.6 表明,将一个 n 阶行列式的行列互换之后,其值不变. 这也就是说,在一个 n 阶行列式中,行和列的地位是平等的,我们只要证明行列式关于它的列具有某种性质,那么它对行也具有相应的性质.

行列式的性质

前面的讨论中已经证明了行列式的一些基本性质,现在对此作一个小结,把这些基本性质罗列如下.

性质 1 行列互换,行列式的值不变,亦即 $|A'|=|A|$.

性质 2 两行(列)互换,行列式值变号.

性质 3 若行列式中某行(列)每个元素分为两个数之和(即某行(列)向量为两向量之和),则该行列式可关于该行(列)拆开成两个行列式之和. 拆开时其它各行(列)均保持不动.

性质 4 行列式中某行(列)有公因子 $\lambda\in K$ 时,λ 可提出行列式外.

性质 5 把行列式的第 j 行(列)加上第 i 行(列)的 k 倍后,其值不变.

性质 6 一个 n 阶方阵 A 不满秩(即 $r(A)<n$)时,其行列式为 0. 特别地,如果 A 有两行(列)元素相同时,$|A|=0$;或 A 有一行(列)元素为 0 时,$|A|=0$.

以上性质 1 即为命题 1.6;性质 3 和 4 是 $\det(A)$ 为行(列)线

性函数的另一种说法;性质 2 和 5 即为命题 1.3;而性质 6 是 det(A)为行列式函数的直接推论.

根据性质 1 和命题 1.3,我们有

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & 0 & & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & & & \\ a_{12} & a_{22} & & 0 \\ \vdots & \vdots & \ddots & \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}.$$

我们已知一个 n 阶方阵可用行、列初等变换化为阶梯形矩阵,从上面的公式可以看出,阶梯形矩阵的行列式值等于其主对角线元素的连乘积. 而性质 2,4,5 说明方阵做初等变换时其行列式值发生什么变化. 因而,这些性质给出计算行列式值的一个有效方法.

例 1 计算下列行列式的值:

$$A = \begin{vmatrix} -2 & 5 & -1 & 3 \\ 1 & -9 & 13 & 7 \\ 3 & -1 & 5 & -5 \\ 2 & 8 & -7 & -10 \end{vmatrix}.$$

解 利用行列式性质 2,4,5 把它化为阶梯形,再利用上面公式计算出它的值,步骤如下:

$$|A| = -\begin{vmatrix} 1 & -9 & 13 & 7 \\ -2 & 5 & -1 & 3 \\ 3 & -1 & 5 & -5 \\ 2 & 8 & -7 & -10 \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & -9 & 13 & 7 \\ 0 & -13 & 25 & 17 \\ 0 & 26 & -34 & -26 \\ 0 & 26 & -33 & -24 \end{vmatrix}$$

$$=-\begin{vmatrix}1 & -9 & 13 & 7\\0 & -13 & 25 & 17\\0 & 0 & 16 & 8\\0 & 0 & 17 & 10\end{vmatrix}=-\begin{vmatrix}1 & -9 & 13 & 7\\0 & -13 & 25 & 17\\0 & 0 & 16 & 8\\0 & 0 & 0 & \frac{3}{2}\end{vmatrix}$$

$$=-(-13)\cdot 16\cdot\frac{3}{2}=312.$$

例 2 计算下列行列式：

$$|A|=\begin{vmatrix}1 & 1 & 1 & 1\\1 & 1 & -1 & -1\\1 & -1 & 1 & -1\\1 & -1 & -1 & 1\end{vmatrix}.$$

解 通过观察发现此行列式第 2,3,4 列元素之和为 0，故利用性质 5，把其第 2,3,4 行的 1 倍加到第 1 行，再利用行列式定义，得

$$|A|=\begin{vmatrix}4 & 0 & 0 & 0\\1 & 1 & -1 & -1\\1 & -1 & 1 & -1\\1 & -1 & -1 & 1\end{vmatrix}=4\begin{vmatrix}1 & -1 & -1\\-1 & 1 & -1\\-1 & -1 & 1\end{vmatrix},$$

然后对右边三阶行列式应用例 1 的方法，得

$$|A|=4\begin{vmatrix}1 & -1 & -1\\0 & 0 & -2\\0 & -2 & 0\end{vmatrix}=-4\begin{vmatrix}1 & -1 & -1\\0 & -2 & 0\\0 & 0 & -2\end{vmatrix}$$

$$=-4\cdot(-2)\cdot(-2)=-16.$$

例 3 计算初等矩阵的行列式.

解 初等矩阵是由单位矩阵做一次初等变换得来的. 已知 $|E|=1$，故由行列式性质 2,4,5 可知

(i) $|P_n(i,j)|=-|E|=-1$.

(ii) $|P_n(\lambda\cdot i)|=\lambda|E|=\lambda\quad(\lambda\neq 0)$.

(iii) $|P_n(k\cdot i,j)|=|E|=1$;

$|P'_n(k\cdot i,j)|=|P_n(k\cdot i,j)|=1$.

行列式对任意行(列)的展开公式

给定 $A=(a_{ij})\in M_n(K)$. 前面用 $A\binom{i}{j}$表示划去 A 的第 i 行第 j 列后所剩的 $n-1$ 阶方阵,其行列式$\left|A\binom{i}{j}\right|$称为 A 中元素 a_{ij}的**余子式**. 为了简单起见,我们把 a_{ij}的余子式简单地写成 M_{ij}(只要从上下文可以清楚地知道它代表的是哪个方阵中元素的余子式,不会产生混淆). 这时 A 的行列式的定义可以写成

$$|A|=\sum_{k=1}^{n}(-1)^{k+1}a_{1k}M_{1k}.$$

它也称为 n 阶行列式$|A|$对其**第一行的展开公式**.

现在我们来证明:行列式可以按它的任意一行或任意一列来展开. 为了把下面所要论证的一般展开公式写的简单明了,我们先介绍一个重要的概念.

定义 设 $A=(a_{ij})$为一 n 阶方阵,M_{ij}为第 i 行第 j 列元素 a_{ij}的余子式. 令

$$A_{ij}=(-1)^{i+j}M_{ij}=(-1)^{i+j}\left|A\binom{i}{j}\right|,$$

称之为元素 a_{ij}的**代数余子式**.

注意余子式 M_{ij}和代数余子式 A_{ij}都是划去 A 的第 i 行第 j 列元素得出的,所以它们的数值恰与 A 的第 i 行元素及第 j 列元素无关. 换句话说,如果改变 A 的第 i 行和第 j 列元素,其它元素不动,那么余子式 M_{ij}和代数余子式 A_{ij}都没有变化. 这一点对我们讨论某些问题是有用的.

例 4 求

$$A=\begin{pmatrix}-1 & 0 & 1\\ 2 & 1 & 0\\ 0 & -3 & 0\end{pmatrix}$$

的全部代数余子式.

解 按定义，有

$$A_{11}=(-1)^{1+1}M_{11}=\begin{vmatrix}1&0\\-3&1\end{vmatrix}=1;$$

$$A_{12}=(-1)^{1+2}M_{12}=-\begin{vmatrix}2&0\\0&1\end{vmatrix}=-2;$$

$$A_{13}=(-1)^{1+3}M_{13}=\begin{vmatrix}2&1\\0&-3\end{vmatrix}=-6;$$

$$A_{21}=(-1)^{2+1}M_{21}=-\begin{vmatrix}0&1\\-3&1\end{vmatrix}=-3;$$

$$A_{22}=(-1)^{2+2}M_{22}=\begin{vmatrix}-1&1\\0&1\end{vmatrix}=-1;$$

$$A_{23}=(-1)^{2+3}M_{23}=-\begin{vmatrix}-1&0\\0&-3\end{vmatrix}=-3;$$

$$A_{31}=(-1)^{3+1}M_{31}=\begin{vmatrix}0&1\\1&0\end{vmatrix}=-1;$$

$$A_{32}=(-1)^{3+2}M_{32}=-\begin{vmatrix}-1&1\\2&0\end{vmatrix}=2;$$

$$A_{33}=(-1)^{3+3}M_{33}=\begin{vmatrix}-1&0\\2&1\end{vmatrix}=-1.$$

利用代数余子式的概念，行列式对第一行的展开公式可以写成

$$|A|=a_{11}A_{11}+a_{12}A_{12}+\cdots+a_{1n}A_{1n}.$$

命题 1.7 n 阶行列式 $|A|$ 可按任意一行和任意一列展开，展开公式为

$$|A|=a_{i1}A_{i1}+a_{i2}A_{i2}+\cdots+a_{in}A_{in}\quad(i=1,2,\cdots,n);$$

$$|A|=a_{1j}A_{1j}+a_{2j}A_{2j}+\cdots+a_{nj}A_{nj}\quad(j=1,2,\cdots,n);$$

证 首先证明对第 i 行的展开公式，再利用行列式性质 1 来证明对任一列的展开公式. $i=1$ 时即为行列式的定义，不需证明. 下面设 $i>1$.

如果把矩阵 A 的 i 行与 $i-1$ 行互换，再与 $i-2$ 行互换，…，最后与第 1 行互换，共经过 $i-1$ 次相邻两行的互换，此时原第 i 行换到第 1 行，而其它行则各向下推移一行，它们之间的相对位置没有变化. 最后得到的矩阵记作 $\overline{A}$. 由行列式性质 2，

$$|\overline{A}| = (-1)^{i-1}|A|,$$

故 $|A|=(-1)^{i-1}|\overline{A}|$. $|A|$ 和 $|\overline{A}|$ 的余子式分别记作 M_{ij} 和 $\overline{M}_{ij}$，它们的代数余子式分别记作 A_{ij} 和 $\overline{A}_{ij}$. 显然有

$$\overline{M}_{1k} = M_{ik} \quad (k = 1,2,\cdots,n).$$

因而

$$\begin{aligned}\overline{A}_{1k} &= (-1)^{1+k}\overline{M}_{1k} = (-1)^{1+k}M_{ik} = (-1)^{1+k}(-1)^{i+k}M_{ik} \\ &= (-1)^{1+i}A_{ik}.\end{aligned}$$

把 $|\overline{A}|$ 按第一行展开（注意 $|\overline{A}|$ 第一行元素为 $a_{i1},a_{i2},\cdots,a_{in}$）

$$\begin{aligned}|\overline{A}| &= a_{i1}\overline{A}_{11} + a_{i2}\overline{A}_{12} + \cdots + a_{in}\overline{A}_{1n} = \sum_{k=1}^{n}a_{ik}\overline{A}_{1k} \\ &= \sum_{k=1}^{n}a_{ik}\cdot(-1)^{i+1}A_{ik} = (-1)^{i+1}\sum_{k=1}^{n}a_{ik}A_{ik}.\end{aligned}$$

而

$$\begin{aligned}|A| &= (-1)^{i-1}|\overline{A}| = (-1)^{i-1}\cdot(-1)^{i+1}\sum_{k=1}^{n}a_{ik}A_{ik} \\ &= \sum_{k=1}^{n}a_{ik}A_{ik}.\end{aligned}$$

下面证明 $|A|$ 对其第 j 列元素的展开公式. A 的第 j 列为 A' 的第 j 行，而 $(A\binom{k}{j})'=A'\binom{j}{k}$，故按行列式性质 1 及 A' 对第 j 行的展开公式，有

$$\begin{aligned}|A| &= |A'| = \sum_{k=1}^{n}(-1)^{j+k}a_{kj}\left|A'\binom{j}{k}\right| \\ &= \sum_{k=1}^{n}a_{kj}(-1)^{j+k}\left|A\binom{k}{j}\right| = \sum_{k=1}^{n}a_{kj}A_{kj}. \ ▌\end{aligned}$$

有了命题 1.7 之后，我们计算行列式时，可以挑其中零最多的

行(列)来展开.

例 5 计算行列式

$$|A| = \begin{vmatrix} -1 & 1 & 1 & 0 & 1 \\ 2 & 1 & 0 & 0 & -1 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 & 2 \end{vmatrix}.$$

解 对第三列展开,得

$$|A| = 1 \cdot A_{13} = (-1)^{1+3} \begin{vmatrix} 2 & 1 & 0 & -1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 2 \end{vmatrix}$$

$$= \begin{vmatrix} 2 & 1 & 0 & -1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 2 \end{vmatrix} = |\overline{A}|.$$

再把最后所得四阶行列式对第二行展开,得

$$|A| = |\overline{A}| = (-1) \cdot A_{22}$$

$$= (-1) \cdot (-1)^{2+2} \begin{vmatrix} 2 & 0 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 2 \end{vmatrix}$$

$$= - \begin{vmatrix} 2 & 0 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 2 \end{vmatrix}$$

$$= -(-1) \cdot (-1)^{2+2} \begin{vmatrix} 2 & -1 \\ 1 & 2 \end{vmatrix} = \begin{vmatrix} 2 & -1 \\ 1 & 2 \end{vmatrix} = 5.$$

最后,我们来介绍一个重要的行列式.

例 6 证明**范德蒙**(Vandermonde)**行列式**

$$|A| = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{1 \leqslant j < i \leqslant n} (a_i - a_j),$$

其中的$\prod$是连乘积的记号. 等式右端表示对 $a_1, a_2, \cdots, a_n$ 这 n 个数作所有可能的差：

$$a_i - a_j \quad (1 \leqslant j < i \leqslant n),$$

然后把它们连乘起来. 具体写出来，就是

$$\begin{aligned} \prod_{1 \leqslant j < i \leqslant n} (a_i - a_j) = (a_2 - a_1)(a_3 - a_1)(a_4 - a_1)\cdots(a_n - a_1) \\ \times (a_3 - a_2)(a_4 - a_2)\cdots(a_n - a_2) \\ \times (a_4 - a_3)\cdots(a_n - a_3) \\ \times \cdots\cdots\cdots\cdots\cdots\cdots \\ \times (a_n - a_{n-1}). \end{aligned}$$

证 采用数学归纳法. 当 $n=2$ 时，有

$$\begin{vmatrix} 1 & 1 \\ a_1 & a_2 \end{vmatrix} = a_2 - a_1,$$

命题成立. 设对 $n-1$ 阶范德蒙行列式命题成立，证明对 n 阶范德蒙行列式 $|A|$，命题也成立.

在 $|A|$ 中将第 n 行减去第 $n-1$ 行的 a_1 倍，第 $n-1$ 行又减去第 $n-2$ 行的 a_1 倍，…. 即由下而上依次把每一行减去它上面一行的 a_1 倍，有

$$|A| = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & a_2 - a_1 & a_3 - a_1 & \cdots & a_n - a_1 \\ 0 & a_2^2 - a_1 a_2 & a_3^2 - a_1 a_3 & \cdots & a_n^2 - a_1 a_n \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & a_2^{n-1} - a_1 a_2^{n-2} & a_3^{n-1} - a_1 a_3^{n-2} & \cdots & a_n^{n-1} - a_1 a_n^{n-2} \end{vmatrix}$$

$$
=\begin{vmatrix} a_2-a_1 & a_3-a_1 & \cdots & a_n-a_1 \\ a_2(a_2-a_1) & a_3(a_3-a_1) & \cdots & a_n(a_n-a_1) \\ \vdots & \vdots & & \vdots \\ a_2^{n-2}(a_2-a_1) & a_3^{n-2}(a_3-a_1) & \cdots & a_n^{n-2}(a_n-a_1) \end{vmatrix}
$$

$$
=(a_2-a_1)(a_3-a_1)\cdots(a_n-a_1)\begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_2 & a_3 & \cdots & a_n \\ a_2^2 & a_3^2 & \cdots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_2^{n-2} & a_3^{n-2} & \cdots & a_n^{n-2} \end{vmatrix}.
$$

最后得到的是一个 $n-1$ 阶范德蒙行列式.根据归纳假设,它等于所有可能的差

$$
(a_i-a_j) \quad (2\leqslant j<i\leqslant n)
$$

的连乘积.而包含 a_1 的差 $a_i-a_1(i=2,3,\cdots,n)$全在前面的因子中出现了,因之,命题对 n 阶范德蒙行列式也成立. ▌

习 题 一

1. 设四阶行列式已经定义,试给出五阶行列式的定义.

2. 给定如下四阶行列式

$$
\begin{vmatrix} -1 & 2 & 0 & 1 \\ 0 & 1 & -1 & 1 \\ -3 & 1 & 0 & 2 \\ -1 & 1 & 0 & 1 \end{vmatrix},
$$

求余子式 M_{13},M_{31},M_{24}.

3. 如果把 n 阶方阵 $A=(a_{ij})$的第 i 行元素换为 $b_1,b_2,\cdots,b_n$,其余不动,得 n 阶方阵 B.问:$|A|$与$|B|$的第 i 行元素的余子式之间有什么关系?

4. 给定 $m\times n$ 矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

把它的第 $i_1, i_2, \cdots, i_r$ 行;第 $j_1, j_2, \cdots, j_r$ 列交叉点处的 r^2 个元素保持其相对位置不变组成一个 r 阶行列式,称为 A 的一个 r **阶子式**.证明:对一个 n 阶方阵 A,如果存在一个正整数 $r \leqslant n$,使 A 的所有 r 阶子式都为零,则

$$|A| = 0.$$

5. 设 A 是奇数阶反对称矩阵:$A' = -A$,证明:$|A| = 0$.

6. 证明

$$\begin{vmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & & \ddots & \ddots & -1 \\ -1 & 0 & \cdots & 0 & 1 \end{vmatrix} = 0.$$

7. 计算下列行列式的值:

(1) $\begin{vmatrix} 1 & 2 & 1 & 4 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{vmatrix}$; (2) $\begin{vmatrix} 1 & -1 & 2 \\ 3 & 2 & 1 \\ 0 & 1 & 4 \end{vmatrix}$;

(3) $\begin{vmatrix} 0 & 1 & 2 & -1 & 4 \\ 2 & 0 & 1 & 2 & 1 \\ -1 & 3 & 5 & 1 & 2 \\ 3 & 3 & 1 & 2 & 1 \\ 2 & 1 & 0 & 3 & 5 \end{vmatrix}$; (4) $\begin{vmatrix} 1 & -2 & 1 & 0 \\ 0 & 3 & -2 & -1 \\ 4 & -1 & 0 & -3 \\ 1 & 2 & -6 & 3 \end{vmatrix}$;

(5) $\begin{vmatrix} 1 & 2 & -1 & 1 \\ 3 & 0 & 1 & 2 \\ 1 & -1 & 2 & 1 \\ 1 & 0 & 3 & -2 \end{vmatrix}$;

(6) $\begin{vmatrix} 1 & \frac{1}{2} & 0 & 1 & -1 \\ 2 & 0 & -1 & 1 & 2 \\ 3 & 2 & 1 & \frac{1}{2} & 0 \\ 1 & -1 & 0 & 1 & 2 \\ 2 & 1 & 3 & 0 & \frac{1}{2} \end{vmatrix}$.

8. 计算下列行列式：

(1) $\begin{vmatrix} x & y & x+y \\ y & x+y & x \\ x+y & x & y \end{vmatrix}$;

(2) $\begin{vmatrix} 1+x & 1 & 1 & 1 \\ 1 & 1-x & 1 & 1 \\ 1 & 1 & 1+y & 1 \\ 1 & 1 & 1 & 1-y \end{vmatrix}$;

(3) $\begin{vmatrix} a^2 & (a+1)^2 & (a+2)^2 & (a+3)^2 \\ b^2 & (b+1)^2 & (b+2)^2 & (b+3)^2 \\ c^2 & (c+1)^2 & (c+2)^2 & (c+3)^2 \\ d^2 & (d+1)^2 & (d+2)^2 & (d+3)^2 \end{vmatrix}$;

(4) $\begin{vmatrix} 3 & 1 & 1 & 1 \\ 1 & 3 & 1 & 1 \\ 1 & 1 & 3 & 1 \\ 1 & 1 & 1 & 3 \end{vmatrix}$;

(5) $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}$.

9. 证明

$$\begin{vmatrix} b+c & c+a & a+b \\ b_1+c_1 & c_1+a_1 & a_1+b_1 \\ b_2+c_2 & c_2+a_2 & a_2+b_2 \end{vmatrix} = 2\begin{vmatrix} a & b & c \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix}.$$

10. 求下列方阵每个元素的代数余子式

$$A = \begin{pmatrix} -3 & 1 & 0 \\ 2 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix},$$

并写出$|A|$对第二行和对第三列的展开公式.

11. 写出行列式$\begin{vmatrix} 4 & -3 & 2 & 1 & -1 \\ 0 & -5 & 3 & 0 & 1 \\ 7 & -1 & 0 & 2 & -1 \\ 1 & 0 & 1 & -1 & -1 \\ 0 & 7 & 3 & 2 & 0 \end{vmatrix}$对第三行的展开公式.

12. 利用数学归纳法证明

$$\begin{vmatrix} a_{11} & \cdots\cdots\cdots & & a_{1n} \\ a_{21} & \cdots & a_{2n-1} & 0 \\ \vdots & \iddots & \iddots & \vdots \\ a_{n1} & 0 & \cdots\cdots & 0 \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} a_{1n}a_{2n-1}\cdots a_{n1}.$$

13. 给定 $n-1$ 个互不相同的数 $a_1,a_2,\cdots,a_{n-1}$,令

$$P(x) = \begin{vmatrix} 1 & x & x^2 & \cdots & x^{n-1} \\ 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_{n-1} & a_{n-1}^2 & \cdots & a_{n-1}^{n-1} \end{vmatrix}.$$

(1) 证明:$P(x)$是一个 $n-1$ 次多项式;

(2) 求出 $P(x)$的 $n-1$ 个根.

14. 设给定 n 阶方阵 $A=(a_{ij})$,将其每个元素 a_{ij}乘以 b^{i-j}($b\neq 0$)后得到 n 阶方阵 B. 证明:$|A|=|B|$.

15. 设给定 n 阶方阵 R 的分块形式

$$R=\begin{pmatrix} A & B \\ 0 & D \end{pmatrix},$$

其中 A 为 k 阶方阵. 证明：$|R|=|A|\cdot|D|$.

§2 行列式理论的应用

在这一节里，我们应用行列式理论来讨论线性方程组和矩阵理论中的若干问题.

齐次线性方程组

首先证明如下命题.

命题 2.1 一个 n 阶方阵 A 满秩的充分必要条件是 $|A|\neq 0$.

证 考虑 A 的初等行、列变换对它的行列式的影响：

(i) 互换 A 的两行(列)，其行列式改变符号；

(ii) 把 A 的某行(列)乘 $c\neq 0$，其行列式增大 c 倍；

(iii) 把 A 的第 j 行(列)加上第 i 行(列)的 k 倍，其行列式之值不变.

因此，A 经过一次初等变换后，其行列式乘上一非零常数. A 可经初等变换化为标准形

$$A\to D=\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}.$$

A 满秩的充分必要条件是 $D=E$. 但 $|A|=k\cdot|D|$(其中 k 为非零常数). 若 $D=E$，则 $|A|=k|E|=k\neq 0$；反之，若 $|A|\neq 0$，则 $|D|=\frac{1}{k}|A|\neq 0$，于是 $D=E$，即 A 满秩. ▌

现在来讨论 n 个未知量 n 个方程的齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0. \end{cases}$$

根据第一章定理 1 的推论，上面的齐次线性方程组有非零解的充分必要条件是其系数矩阵 A 的秩 $r(A)$ 小于未知量个数 n. 现在 A 是一个 n 阶方阵，由命题 2.1，$r(A)<n$ 的充分必要条件是 $|A|=0$. 故有如下重要结论.

定理 1 n 个未知量 n 个方程的齐次线性方程组有非零解的充分必要条件是其系数矩阵的行列式为零.

这个定理告诉我们：如果这样的齐次线性方程组系数矩阵不为零时，它就只有零解. 注意这定理仅限于 n 个未知量 n 个方程的情况.

逆 矩 阵

首先我们介绍数学中一个常用的记号. 命

$$\delta_{ij} = \begin{cases} 1, & \text{当 } i = j \text{ 时}; \\ 0, & \text{当 } i \neq j \text{ 时}. \end{cases}$$

这个记号称为**克朗涅克**(Kronecker)**记号**. 利用这个记号，许多数学式子就可以写的比较简洁. 例如，单位矩阵 E 可以写成 (δ_{ij})，即 E 的 i 行 j 列元素为 δ_{ij}.

命题 2.2 n 阶方阵 $A=(a_{ij})$ 的行列式 $|A|$ 和它的代数余子式有如下关系

$$\begin{aligned} a_{i1}A_{j1}+a_{i2}A_{j2}+\cdots+a_{in}A_{jn}=\delta_{ij}|A|, \\ a_{1i}A_{1j}+a_{2i}A_{2j}+\cdots+a_{ni}A_{nj}=\delta_{ij}|A|, \end{aligned} \quad (i,j=1,2,\cdots,n).$$

证 当 $i=j$ 时上面两个等式即为命题 1.7. 下面来证 $i\neq j$ 的情况. 我们只要证明了第一个关于行的公式，那么，由行列式的性质 1，第二个关于列的公式就随之成立.

把 A 的第 j 行元素换成 $a_{i1},a_{i2},\cdots,a_{in}$，得矩阵 $\overline{A}$. $\overline{A}$ 的 i,j 两行相同，故 $|\overline{A}|=0$. 把 $|\overline{A}|$ 对第 j 行展开，就有

$$a_{i1}A_{j1}+a_{i2}A_{j2}+\cdots+a_{in}A_{jn}=|\overline{A}|=0.$$

其中利用了 $\overline{A}$ 与 A 仅是第 j 行不相同，故 $|\overline{A}|$ 第 j 行元素的代数余子式与 $|A|$ 的第 j 行元素的代数余子式相同. ▌

给定一个 n 阶方阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

用 $|A|$ 的代数余子式排成如下一个 n 阶方阵

$$A^*=\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$$

（即 A^* 的 i 行 j 列处放置代数余子式 A_{ji}），A^* 称为 A 的**伴随矩阵**. 根据命题 2.2，有

$$AA^*=\left(\sum_{k=1}^{n}a_{ik}A_{jk}\right)=(\delta_{ij}|A|)=|A|\cdot E,$$

$$A^*A=\left(\sum_{k=1}^{n}A_{ki}a_{kj}\right)=(\delta_{ij}|A|)=|A|\cdot E.$$

如果 $|A|\neq 0$，则有

$$\left(\frac{1}{|A|}A^*\right)A=A\left(\frac{1}{|A|}A^*\right)=E.$$

这表示此时 A 可逆，且

$$A^{-1}=\frac{1}{|A|}A^*.$$

反之，若 A 可逆，则由第二章定理 2 知 A 满秩，再由命题 2.1 知 $|A|\neq 0$，于是 $\frac{1}{|A|}A^*$ 有意义，那么，它就是 A 的逆矩阵. 故有

定理 2 n 阶方阵 A 可逆的充分必要条件是 $|A|\neq 0$. 在 A 可逆时,有

$$A^{-1}=\frac{1}{|A|}A^{*}.$$

例 1 给定矩阵

$$A=\begin{pmatrix}-1 & 0 & 1\\ 2 & 1 & 0\\ 0 & -3 & 1\end{pmatrix},$$

有

$$|A|=\begin{vmatrix}-1 & 0 & 1\\ 2 & 1 & 0\\ 0 & -3 & 1\end{vmatrix}=\begin{vmatrix}-1 & 0 & 0\\ 2 & 1 & 2\\ 0 & -3 & 1\end{vmatrix}$$

$$=-\begin{vmatrix}1 & 2\\ -3 & 1\end{vmatrix}=-7\neq 0.$$

故 A 可逆. 在 §1 的例 4 中已经算出 $|A|$ 的全部代数余子式,于是可写出 A 的伴随矩阵

$$A^{*}=\begin{pmatrix}A_{11} & A_{21} & A_{31}\\ A_{12} & A_{22} & A_{32}\\ A_{13} & A_{23} & A_{33}\end{pmatrix}=\begin{pmatrix}1 & -3 & -1\\ -2 & -1 & 2\\ -6 & -3 & -1\end{pmatrix}.$$

那么,A 的逆矩阵是

$$A^{-1}=\frac{1}{|A|}A^{*}=-\frac{1}{7}\begin{pmatrix}1 & -3 & -1\\ -2 & -1 & 2\\ -6 & -3 & -1\end{pmatrix}$$

$$=\begin{pmatrix}-\frac{1}{7} & \frac{3}{7} & \frac{1}{7}\\ \frac{2}{7} & \frac{1}{7} & -\frac{2}{7}\\ \frac{6}{7} & \frac{3}{7} & \frac{1}{7}\end{pmatrix}.$$

下面再来讨论 n 个未知量 n 个方程的线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n. \end{cases} \tag{1}$$

命

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

则方程组可写成

$$AX = B. \tag{2}$$

现在(1)的系数矩阵 A 是一个 n 阶方阵. 我们有如下两个结论：

1. 方程组(1)有唯一解的充分必要条件是 $|A| \neq 0$.

(i) 若 $|A| \neq 0$, 由命题 2.1 知, $r(A) = n$. 从而方程组(1)的增广矩阵 $\overline{A}$ 的秩 $r(\overline{A}) = n = r(A)$(因 $\overline{A}$ 只有 n 行, 其秩不超过 n, 而 $r(\overline{A}) \geqslant r(A) = n$), 故方程组有解. 再由第一章定理 3 的(i)知解唯一.

(ii) 若方程组有唯一解, 则由第一章定理 3 的(ii)可知, $r(A) = n$. 再由命题 2.1 知, $|A| \neq 0$.

2. 当 $|A| \neq 0$, 方程组(1)有唯一解时, 命

$$X = A^{-1}B. \tag{3}$$

代入(2)式, 即知(3)式就是方程组的唯一解.

现在用伴随矩阵表示 A^{-1}, 代入(3)式, 得

$$X = A^{-1}B = \frac{1}{|A|}A^*B$$

$$= \frac{1}{|A|}\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$$= \frac{1}{|A|}\begin{pmatrix} \sum_{k=1}^{n} A_{k1}b_k \\ \sum_{k=1}^{n} A_{k2}b_k \\ \vdots \\ \sum_{k=1}^{n} A_{kn}b_k \end{pmatrix}.$$

命

$$|A_i| = \sum_{k=1}^{n} A_{ki}b_k$$

$$= \begin{vmatrix} a_{11} & \cdots & a_{1\,i-1} & b_1 & a_{1\,i+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2\,i-1} & b_2 & a_{2\,i+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n\,i-1} & b_n & a_{n\,i+1} & \cdots & a_{nn} \end{vmatrix}.$$

$|A_i|$ 恰为把 $|A|$ 的第 i 列换成方程组(1)的常数项而得的 n 阶行列式. 此时方程组(1)的解可表为

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{|A|}\begin{pmatrix} |A_1| \\ |A_2| \\ \vdots \\ |A_n| \end{pmatrix}.$$

由此即得如下重要结论:

定理 3 n 个未知量 n 个方程的线性方程组(1)的系数矩阵的行列式 $|A| \neq 0$ 时,它有唯一的一组解

$$x_1 = \frac{|A_1|}{|A|}, \quad x_2 = \frac{|A_2|}{|A|}, \quad \cdots, \quad x_n = \frac{|A_n|}{|A|},$$

其中 $|A_i|$ $(i=1,2,\cdots,n)$ 是把 $|A|$ 的第 i 列换成方程组的常数项而得的 n 阶行列式.

定理 3 通常称为**克莱姆**(Cramer)**法则**.

例 2 解方程组

$$\begin{cases} 2x_1+\ x_2-5x_3+\ x_4=8, \\ x_1-3x_2 \qquad\quad -6x_4=9, \\ \qquad 2x_2-\ x_3+2x_4=-5, \\ x_1+4x_2-7x_3+6x_4=0. \end{cases}$$

解 先算系数矩阵的行列式

$$|A|=\begin{vmatrix} 1 & 1 & -5 & 1 \\ 2 & -3 & 0 & -6 \\ 0 & 2 & -1 & 2 \\ 1 & 4 & -7 & 6 \end{vmatrix}=27\neq 0,$$

故克莱姆法则可以应用. 由于

$$|A_1|=\begin{vmatrix} 8 & 1 & -5 & 1 \\ 9 & -3 & 0 & -6 \\ -5 & 2 & -1 & 2 \\ 0 & 4 & -7 & 6 \end{vmatrix}=81,$$

$$|A_2|=\begin{vmatrix} 2 & 8 & -5 & 1 \\ 1 & 9 & 0 & -6 \\ 0 & -5 & -1 & 2 \\ 1 & 0 & -7 & 6 \end{vmatrix}=-108,$$

$$|A_3|=\begin{vmatrix} 2 & 1 & 8 & 1 \\ 1 & -3 & 9 & -6 \\ 0 & 2 & -5 & 2 \\ 1 & 4 & 0 & 6 \end{vmatrix}=-27,$$

$$|A_4|=\begin{vmatrix} 2 & 1 & -5 & 8 \\ 1 & -3 & 0 & 9 \\ 0 & 2 & -1 & -5 \\ 1 & 4 & -7 & 0 \end{vmatrix}=27,$$

故方程组的唯一的一组解为

$$x_1 = \frac{|A_1|}{|A|} = 3, \quad x_2 = \frac{|A_2|}{|A|} = -4,$$

$$x_3 = \frac{|A_3|}{|A|} = -1, \quad x_4 = \frac{|A_4|}{|A|} = 1.$$

从上面的例子可以看出，如用克莱姆法则去解线性方程组，其计算量是很大的，远不如用第一章的矩阵消元法简单. 但这个法则对理论上讨论某些问题是有用的. 还应当注意：克莱姆法则只能用于 n 个未知量 n 个方程且系数矩阵的行列式不为零的线性方程组.

矩阵乘积的行列式

现在来考虑两个 n 阶方阵乘积的行列式.

命题 2.3 设 P 是 n 阶初等矩阵，B 是任一 n 阶方阵，则

$$|PB| = |P| \cdot |B|.$$

证 PB 相当于对 B 作一次初等行变换. 下面分三种情况讨论：

(i) $P=P_n(i,j)$，则 $|P|=-1$. 而 PB 是互换 B 的 i,j 两行，由行列式性质 2，有 $|PB|=-|B|=|P|\cdot|B|$；

(ii) $P=P_n(c\cdot i)$，则 $|P|=c$. 而 PB 是把 B 的第 i 行乘以 c，由行列式性质 3，有 $|PB|=c|B|=|P|\cdot|B|$；

(iii) $P=P_n(k\cdot i,j)$，则 $|P|=1$. 而 PB 是把 B 的第 j 行加上第 i 行的 k 倍，由行列式性质 7，$|PB|=|B|=|P|\cdot|B|$. ▌

推论 设 $P_1,P_2,\cdots,P_s$ 是 n 阶初等矩阵，B 是任一方阵，则有

$$|P_1P_2\cdots P_sB| = |P_1|\cdot|P_2|\cdots|P_s|\cdot|B|.$$

证 反复利用命题 2.3，有

$$\begin{aligned}|P_1P_2\cdots P_sB| &= |P_1|\cdot|P_2\cdots P_sB| = \cdots \\ &= |P_1|\cdot|P_2|\cdots|P_s|\cdot|B|. \ ▌\end{aligned}$$

定理 4 对任意两个 n 阶方阵 A,B 有

$$|AB| = |A|\cdot|B|.$$

证 分两种情况讨论.

(i) 若$|A|=0$,则$r(A)<n$. 由第二章定理1,有

$$r(AB)\leqslant \min\{r(A),r(B)\}<n.$$

故$|AB|=0=|A|\cdot|B|$.

(ii) 若$|A|\neq 0$,则由命题2.1,$r(A)=n$. 再由第二章命题2.2,A可表为初等矩阵的乘积:

$$A=P_1P_2\cdots P_s.$$

于是由命题2.3的推论,有

$$|A|=|P_1|\cdot|P_2|\cdots|P_s|,$$

$$\begin{aligned}|AB|&=|P_1P_2\cdots P_sB|=|P_1|\cdot|P_2|\cdots|P_s|\cdot|B|\\&=|A|\cdot|B|.\end{aligned}$$ ▌

习 题 二

1. 设$\alpha_1,\alpha_2,\cdots,\alpha_s$是一个线性无关向量组,而

$$\beta_i=\sum_{j=1}^{s}a_{ij}\alpha_j,\quad i=1,2,\cdots,s.$$

证明$\beta_1,\beta_2,\cdots,\beta_s$线性无关的充分必要条件是下面的$s$阶行列式

$$\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1s}\\a_{21}&a_{22}&\cdots&a_{2s}\\\vdots&\vdots&&\vdots\\a_{s1}&a_{s2}&\cdots&a_{ss}\end{vmatrix}\neq 0.$$

2. 给定$m\times n$矩阵

$$A\doteq\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix},$$

证明:若$r(A)=m$,则A有一个m阶子式不为零.

(提示:考虑A的列向量组的极大线性无关部分组.)

3. 给定$m\times n$矩阵A,证明:当A有一个m阶子式不为零

时，$r(A)=m$.

4. 给定 $m\times n$ 矩阵 A，证明：$r(A)=r$ 的充分必要条件是：A 有一个 r 阶子式不为零，而所有 $r+1$ 阶子式全为零.

5. 设 $t_1,t_2,\cdots,t_r$ 是互不相同的数，$r\leqslant n$. 证明：

$$\alpha_i=(1,t_i,\cdots,t_i^{n-1})\quad(i=1,2,\cdots,r)$$

是线性无关的.

6. 给定齐次线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+a_{13}x_3=\lambda x_1,\\a_{21}x_1+a_{22}x_2+a_{23}x_3=\lambda x_2,\\a_{31}x_1+a_{32}x_2+a_{33}x_3=\lambda x_3,\end{cases}$$

问 λ 满足什么条件时它才有（且一定有）非零解？

7. 给定 n 阶方阵 $A=(a_{ij})$，证明：若

$$|a_{ii}|>\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}|\quad(i=1,2,\cdots,n),$$

则 $|A|\neq 0$.

（提示：用反证法. 若 $|A|=0$，则齐次线性方程组 $AX=0$ 有非零解

$$x_1=k_1,\ x_2=k_2,\ \cdots,\ x_n=k_n.$$

考虑诸 k_i 中绝对值最大者.）

8. 给定矩阵

$$A=\begin{pmatrix}-1&2&0\\-3&1&1\\0&2&0\end{pmatrix},$$

求它的伴随矩阵 A^*.

9. 利用伴随矩阵求下列矩阵的逆矩阵：

(1) $A=\begin{pmatrix}1&1&-1\\2&1&0\\1&-1&0\end{pmatrix}$； (2) $A=\begin{pmatrix}2&2&3\\1&-1&0\\-1&2&1\end{pmatrix}$；

(3) $A=\begin{pmatrix}1 & 0 & -1 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & -1 & 1\\ 0 & 0 & 0 & -1\end{pmatrix}$.

10. 给定线性方程组

$$\begin{cases} a_{11}x_1+ a_{12}x_2+\cdots+ a_{1n}x_n=0,\\ a_{21}x_1+ a_{22}x_2+\cdots+ a_{2n}x_n=0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ a_{n-1\,1}x_1+a_{n-1\,2}x_2+\cdots+a_{n-1\,n}x_n=0,\end{cases}$$

以 M_i 表其系数矩阵划去第 i 列后所剩 $n-1$ 阶方阵的行列式. 证明:

(i) $(M_1,-M_2,\cdots,(-1)^{n-1}M_n)$是方程组的解;

(ii) 若方程组系数矩阵的秩为 $n-1$,则方程组的解全是$(M_1,-M_2,\cdots,(-1)^{n-1}M_n)$的倍数.

11. 证明: 对 n 阶方阵 A,有 $|A^*|=|A|^{n-1}$.

12. 设 A 是 n 阶方阵,$n\geqslant 2$. 证明:

$$r(A^*)=\begin{cases} n, & 当\ r(A)=n,\\ 1, & 当\ r(A)=n-1,\\ 0, & 当\ r(A)<n-1.\end{cases}$$

13. 设 A,B,T 均为 n 阶实数方阵,T 可逆. 证明:

(1) 若 $B=T^{-1}AT$,则 $|B|=|A|$;

(2) 若 $B=T'AT$,且 $|A|>0$,则 $|B|>0$.

14. 设将 n 阶方阵 R 分块

$$R=\begin{pmatrix}A & B\\ C & D\end{pmatrix},$$

其中 A 为 k 阶可逆方阵. 证明:

$$|R|=|A|\cdot|D-CA^{-1}B|.$$

15. 在平面直角坐标系 OXY 内给定二次曲线

$$a_{11}x^2+2a_{12}xy+a_{22}y^2+2b_1x+2b_2y+c=0,$$

证明:

$$F=\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{12} & a_{22} & b_2 \\ b_1 & b_2 & c \end{vmatrix}$$

是坐标变换

$$\begin{cases} x=\cos\theta\cdot x'-\sin\theta\cdot y'+x_0, \\ y=\sin\theta\cdot x'-\cos\theta\cdot y'+y_0 \end{cases}$$

的不变量.

(提示:将上述二次多项式写成

$$(x \quad y \quad 1)\begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{12} & a_{22} & b_2 \\ b_1 & b_2 & c \end{pmatrix}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix},$$

再把坐标变换公式用矩阵表示,然后代入计算.)

16. 利用克莱姆法则解下列线性方程组:

(1) $$\begin{cases} 2x_1-x_2+3x_3+2x_4=6, \\ 3x_1-3x_2+3x_3+2x_4=5, \\ 3x_1-x_2-x_3+2x_4=3, \\ 3x_1-x_2+3x_3-x_4=4. \end{cases}$$

(2) $$\begin{cases} x_1+2x_2+3x_3-2x_4=6, \\ 2x_1-x_2-2x_3-3x_4=8, \\ 3x_1+2x_2-x_3+2x_4=4, \\ 2x_1-3x_2+3x_3+x_4=-8. \end{cases}$$

17. 给定 n 个互不相同的数 $a_1,a_2,\cdots,a_n$,又任意给定 n 个数 $b_1,b_2,\cdots,b_n$. 证明:存在一个次数小于 n 的多项式 $f(x)$,使

$$f(a_i)=b_i \quad (i=1,2,\cdots,n),$$

且这样的多项式是唯一的.

* § 3. 行列式的完全展开式

按照行列式的定义,一个 n 阶行列式 $|A|$ 可展开为

$$|A| = a_{11}M_{11} - a_{12}M_{12} + \cdots + (-1)^{n+1}a_{1n}M_{1n},$$

其中 $M_{11}, M_{12}, \cdots, M_{1n}$ 都是 $n-1$ 阶行列式，它们又可以按第一行展开，用 $n-2$ 阶行列式来表示. …，继续这个工作，最后就得到 n 阶行列式的完全展开式，也就是利用行列式中的 n^2 个元素直接表示行列式数值的一般公式. 在这一节里，我们来导出这个公式.

首先我们给出排列的一些基本概念.

给定 n 个互不相同的自然数，把它们按一定次序排列起来：

$$i_1 i_2 \cdots i_n,$$

称为该 n 个自然数的一个排列. 在上述排列中，如果有一个较大的自然数排在一个较小的自然数前面，则称为一个**反序**. 例如，2，3，5，7 这四个自然数的一个排列 7325，其中 3 在 2 前，是一个反序；7 在 2 前，是一个反序；7 在 3 前，是一个反序；7 在 5 前，也是一反序，故此排列共有 4 个反序.

一个排列中包含的反序的总数称为该排列的**反序数**. 排列 $i_1 i_2 \cdots i_n$ 的反序数记作 $N(i_1 i_2 \cdots i_n)$. 例如我们有 $N(7325)=4$. 一个排列的反序数是奇数时，该排列称为**奇排列**；如果反序数是偶数，则称为**偶排列**. 例如，7325 是一个偶排列，而因为 $N(7235)=3$，故 7235 是一个奇排列.

下面开始讨论行列式的完全展开式. 我们先来分析一下二、三阶行列式的展开式. 二阶行列式的展开式是

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

它是由两项组成的，每项是行列式中属于不同行和不同列的两个元素的乘积. 如果把行角标按自然数的顺序排列，它的一般项可表为

$$a_{1i_1} a_{2i_2},$$

而 $i_1 i_2$ 是自然数 1，2 的一个排列. 这样的排列一共有两个：12 和 21. 而展开式中恰好有两项：$N(12)=0$ 是偶数，相应的项 $a_{11}a_{12}$ 前

面取正号；$N(21)=1$ 是奇数，相应的项 $a_{12}a_{21}$ 前面取负号.

再看三阶行列式，其完全展开式是

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31}$$

$$- a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

其中蕴涵着如下几条规律：

(i) 展开式由 $6=3!$ 项组成；

(ii) 每一项是行列式中不同行又不同列的三个元素的乘积. 如把其行角标按自然顺序排列，则其一般形式为

$$a_{1i_1}a_{2i_2}a_{3i_3},$$

其中 $i_1i_2i_3$ 是自然数 1，2，3 的一个排列，这种排列一共有 $3!=6$ 个，它们恰好对应于三阶行列式的完全展开式中包含的 6 项；

(iii) 每项前面带有正负号. 如 $i_1i_2i_3$ 为偶排列，则 $a_{1i_1}a_{2i_2}a_{3i_3}$ 这一项前面带正号，如 $i_1i_2i_3$ 为奇排列，则该项前面带负号.

把上面的分析综合起来，得出下面的公式

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$= \sum_{(i_1i_2i_3)} (-1)^{N(i_1i_2i_3)} \cdot a_{1i_1}a_{2i_2}a_{3i_3},$$

其中 $i_1i_2i_3$ 是自然数 1，2，3 的一个排列，和号 $\sum$ 表示对所有可能的这种排列求和.

上面之所以利用 1，2，3 这三个自然数的排列来表达三阶行列式的完全展开式，是因为行列式中元素的列角标用 1，2，3 来表示的缘故. 如果行列式中元素的列角标改用其它自然数来表示（这也是可以的），那么，我们也就要相应地用其它自然数的排列来表示三阶行列式的完全展开式了.

现在我们可以对 n 阶行列式的完全展开式作如下的推测：

(i) 它应由 $n!$ 项组成；

(ii) 每项是行列式中 n 个不同行且不同列的元素的乘积. 如把它们按行角标的自然顺序排列，则其一般形式为

$$a_{1i_1}a_{2i_2}a_{ni_n},$$

其中 $i_1i_2\cdots i_n$ 是 $1,2,\cdots,n$ 这 n 个自然数的一个排列(如果行列式的列角标不用 $1,2,\cdots,n$ 表示，而用其它自然数表示，那就改用其它自然数的排列). 这样的排列共有 $n!$ 个，就对应 n 阶行列式的完全展开式中应包含的 $n!$ 项；

(iii) 每项前面应带正、负号. 如果 $i_1i_2\cdots i_n$ 是一个偶排列，则 $a_{1i_1}a_{2i_2}\cdots a_{ni_n}$ 这一项前面带正号；而如果 $i_1i_2\cdots i_n$ 是奇排列，则该项前面带负号.

下面证明这个推测.

定理 5 n 阶行列式有如下完全展开式

$$|A|=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$=\sum_{(i_1i_2\cdots i_n)}(-1)^{N(i_1i_2\cdots i_n)}\cdot a_{1i_1}a_{2i_2}\cdots a_{ni_n},$$

其中 $i_1i_2\cdots i_n$ 是前 n 个自然数 $1,2,\cdots,n$ 的一个排列，和号 $\sum$ 表示对所有可能的这种排列求和.

证 采用数学归纳法. $n=2$ 时，从前面的分析已知命题正确. 设对 $n-1$ 阶行列式命题成立，证明对 n 阶行列式命题也成立.

按行列式的定义，有

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$=a_{11}M_{11}-a_{12}M_{12}+\cdots+(-1)^{n+1}a_{1n}M_{1n}.$$

现在我们取出其中第 k 项：$(-1)^{1+k}a_{1k}M_{1k}$进行分析. 按归纳假设，有

$$M_{1k}=\begin{vmatrix} a_{21} & \cdots & a_{2\,k-1} & a_{2\,k+1} & \cdots & a_{2n} \\ a_{31} & \cdots & a_{3\,k-1} & a_{3\,k+1} & \cdots & a_{3n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n\,k-1} & a_{n\,k+1} & \cdots & a_{nn} \end{vmatrix}$$

$$=\sum_{(i_2i_3\cdots i_n)}(-1)^{N(i_2i_3\cdots i_n)}\cdot a_{2i_2}a_{3i_3}\cdots a_{ni_n},$$

其中 $i_2i_3\cdots i_n$ 是自然数 $1,2,\cdots,k-1,k+1,\cdots,n$ 的一个排列，而和号$\sum$则表示对所有可能的这种排列求和. 这个展开式的写法和一般 $n-1$ 阶行列式的完全展开式的写法有些不同. 这原因是 M_{1k}中元素的行角标用的是自然数 $2,3,\cdots,n$，而不是 $1,2,\cdots,n-1$，所以后面展开式第一个角标的排列是 $2,3,\cdots,n$；M_{1k}中元素的列角标是自然数 $1,2,\cdots,k-1,k+1,\cdots,n$，而不是 $1,2,\cdots,n-1$，所以现在求和式中第二个角标所成的排列 $i_2i_3\cdots i_n$ 是 $1,2,\cdots,k-1,k+1,\cdots,n$ 的一个排列. 这只是写法上的不同，对问题的实质没有影响.

于是，我们有

$$\begin{aligned} & a_{1k}\cdot(-1)^{1+k}M_{1k} \\ &= a_{1k}\cdot(-1)^{1+k}\sum_{(i_2i_3\cdots i_n)}(-1)^{N(i_2i_3\cdots i_n)}\cdot a_{2i_2}a_{3i_3}\cdots a_{ni_n} \\ &= \sum_{(i_2i_3\cdots i_n)}(-1)^{(k-1)+N(i_2i_3\cdots i_n)}\cdot a_{1k}a_{2i_2}a_{3i_3}\cdots a_{ni_n}, \end{aligned}$$

其中用到了：$(-1)^{1+k}=(-1)^{k-1}$. 现在考察排列

$$ki_2i_3\cdots i_n.$$

这是 $1,2,\cdots,n$ 这 n 个自然数的一个排列，其中有 $k-1$ 个比 k 小的自然数(即 $1,2,\cdots,k-1$)排在 k 的后面，故此排列关于 k 有 $k-1$个反序，其它反序与 k 无关. 从而有

$$N(ki_2i_3\cdots i_n)=(k-1)+N(i_2i_3\cdots i_n).$$

代入(1)式，得

$$a_{1k}\cdot(-1)^{1+k}M_{1k}=\sum_{(i_2i_3\cdots i_n)}(-1)^{N(ki_2i_3\cdots i_n)}\cdot a_{1k}a_{2i_2}a_{3i_3}\cdots a_{ni_n}. \tag{2}$$

现在考察 $1,2,\cdots,n$ 这 n 个文字的所有可能的排列

$$i_1i_2\cdots i_n.$$

当 i_1 取自然数 k 时，$i_2i_3\cdots i_n$ 取 $1,2,\cdots,k-1,k+1,\cdots,n$ 的所有可能的排列. 再让 k 取遍 $1,2,\cdots,n$，这就得出 $1,2,\cdots,n$，这 n 个文字的所有可能的排列，而且其中没有重复. 根据这一分析，如把(2)式所得结果令 $k=1,2,\cdots,n$ 代入再相加，即得

$$\begin{aligned}|A|&=\sum_{k=1}^{n}a_{1k}\cdot(-1)^{1+k}M_{1k}\\&=\sum_{k=1}^{n}\sum_{(i_2i_3\cdots i_n)}(-1)^{N(ki_2i_3\cdots i_n)}\cdot a_{1k}a_{2i_2}\cdots a_{ni_n}\\&=\sum_{(i_1i_2\cdots i_n)}(-1)^{N(i_1i_2\cdots i_n)}\cdot a_{1i_1}a_{2i_2}\cdots a_{ni_n}. \blacksquare\end{aligned}$$

例 1 计算行列式

$$|A|=\begin{vmatrix}a_{11} & \cdots\cdots & a_{1n}\\ a_{21} & \cdots\; a_{2\,n-1} & 0\\ \vdots & \ddots & \vdots\\ a_{n1} & 0\;\cdots & 0\end{vmatrix}.$$

解 按定理 5，有

$$|A|=\sum_{(i_1i_2\cdots i_n)}(-1)^{N(i_1i_2\cdots i_n)}\cdot a_{1i_1}a_{2i_2}\cdots a_{ni_n}.$$

只需考察展开式中可能不为零的项. 显然，此时必有 $i_n=1$(第 n 列除第一个元素不为零外其它元素均为零)，$i_{n-1}=2,\cdots,i_2=n-1,i_1=n$. 故

$$|A|=(-1)^{N(n(n-1)\cdots 2\cdot 1)}\cdot a_{1n}a_{2\,n-1}\cdots a_{n1}.$$

因为 $N(n(n-1)\cdots 21)=1+2+\cdots+(n-1)=n(n-1)/2$，故

$$|A|=(-1)^{\frac{n(n-1)}{2}}\cdot a_{1n}a_{2\,n-1}\cdots a_{n1}.$$

例 2 设 $a_{ij}(t)$是 t 的可微函数,求证:

$$\frac{\mathrm{d}|A|}{\mathrm{d}t}=\frac{\mathrm{d}}{\mathrm{d}t}\begin{vmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t) \\ \vdots & \vdots & & \vdots \\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nn}(t) \end{vmatrix}$$

$$=\sum_{k=1}^{n}\begin{vmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ \vdots & \vdots & & \vdots \\ a'_{k1}(t) & a'_{k2}(t) & \cdots & a'_{kn}(t) \\ \vdots & \vdots & & \vdots \\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nn}(t) \end{vmatrix}.$$

证 利用定理 5,有

$$\frac{\mathrm{d}|A|}{\mathrm{d}t}=\frac{\mathrm{d}}{\mathrm{d}t}\sum_{(i_1i_2\cdots i_n)}(-1)^{N(i_1i_2\cdots i_n)}\cdot a_{1i_1}(t)a_{2i_2}(t)\cdots a_{ni_n}(t)$$

$$=\sum_{k=1}^{n}\sum_{(i_1i_2\cdots i_n)}(-1)^{N(i_1i_2\cdots i_n)}a_{1i_1}(t)\cdots a'_{ki_k}(t)\cdots a_{ni_n}(t)$$

$$=\sum_{k=1}^{n}\begin{vmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ \vdots & \vdots & & \vdots \\ a'_{k1}(t) & a'_{k2}(t) & \cdots & a'_{kn}(t) \\ \vdots & \vdots & & \vdots \\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nn}(t) \end{vmatrix}.$$

现在利用定理 5 证明如下命题.

命题 3.1 一个 n 阶方阵 A 有如下分块形式

$$A=\begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \end{pmatrix},$$

其中 A_1 是 r 阶方阵,A_2 是 $n-r$ 阶方阵.则

$$|A|=|A_1|\cdot|A_2|.$$

证 把 A 具体写出来

$$A=\begin{bmatrix} a_{11} & \cdots & a_{1r} & a_{1\,r+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} & a_{r\,r+1} & \cdots & a_{rn} \\ 0 & \cdots & 0 & a_{r+1\,r+1} & \cdots & a_{r+1\,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & a_{n\,r+1} & \cdots & a_{nn} \end{bmatrix}.$$

按定理 5，有

$$|A| = \sum_{(i_1 i_2 \cdots i_n)} (-1)^{N(i_1 i_2 \cdots i_n)} a_{1i_1} a_{2i_2} \cdots a_{ri_r} a_{r+1\,i_{r+1}} \cdots a_{ni_n}.$$

因为 A 的 $r+1,\cdots,n$ 行中前 r 个元素全为零，故 $|A|$ 的上述展开式中，$i_{r+1},\cdots,i_n$ 仅能在 $r+1,\cdots,n$ 这 $n-r$ 个自然数中取，即是它们的一个排列. 从而，$i_1,\cdots,i_r$ 应是 $1,2,\cdots,r$ 这 r 个自然数的一个排列. 此时显然有

$$N(i_1 \cdots i_r i_{r+1} \cdots i_n) = N(i_1 \cdots i_r) + N(i_{r+1} \cdots i_n),$$

故

$$\begin{aligned} |A| &= \left(\sum_{(i_1 \cdots i_r)} (-1)^{N(i_1 \cdots i_r)} \cdot a_{1i_1} \cdots a_{ri_r} \right) \\ &\quad \times \left(\sum_{(i_{r+1} \cdots i_n)} (-1)^{N(i_{r+1} \cdots i_n)} \cdot a_{r+1\,i_{r+1}} \cdots a_{ni_n} \right) \\ &= |A_1| \cdot |A_2|. \ \blacksquare \end{aligned}$$

特别地，如果 A 是准对角形

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix},$$

则 $$|A| = |A_1| \cdot |A_2|.$$

利用数学归纳法，不难得出

推论 设 A 是准对角矩阵

$$A = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{bmatrix},$$

则

$$|A| = |A_1||A_2|\cdots|A_s|.$$

习 题 三

1. 求下列排列的反序数,并判断它是奇排列还是偶排列.

23145; 985467321; 375149;

$n(n-1)(n-2)\cdots321$; $(2n+1)(2n-1)\cdots531$.

2. 选择 i 与 $k(1\leqslant i,k\leqslant 9)$使

(1) $1274i56k9$ 成偶排列;

(2) $1i25k4897$ 成奇排列.

3. 在六阶行列式中,

$$a_{23}a_{31}a_{42}a_{56}a_{14}a_{65}$$

以及

$$a_{32}a_{43}a_{14}a_{51}a_{66}a_{25}$$

这两项应带什么符号?

4. 写出四阶行列式中所有带负号,且包含因子 a_{23}的项.

5. 证明:

$$\begin{vmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \\ b_1 & b_2 & b_3 & b_4 & b_5 \\ c_1 & c_2 & 0 & 0 & 0 \\ d_1 & d_2 & 0 & 0 & 0 \\ e_1 & e_2 & 0 & 0 & 0 \end{vmatrix} = 0.$$

6. 求

$$f(x) = \begin{vmatrix} 2x & x & 1 & 2 \\ 1 & x & 1 & -1 \\ 3 & 2 & x & 1 \\ 1 & 1 & 1 & x \end{vmatrix}$$

中 x^4 与 x^3 的系数.

7. 由

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{vmatrix} = 0$$

证明：前 n 个自然数 $1,2,\cdots,n$ 所组成的排列中，奇、偶排列各占一半.

8. 求

$$\sum_{j_1j_2\cdots j_n} \begin{vmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_n} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_n} \\ \vdots & \vdots & & \vdots \\ a_{nj_1} & a_{nj_2} & \cdots & a_{nj_n} \end{vmatrix},$$

其中和号表示对前 n 个自然数的所有可能的排列 $j_1j_2\cdots j_n$ 求和.

9. 设 $A=(a_{ij})$是一个 n 阶方阵，证明：

$$|\lambda E - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

是 λ 的首项系数为 1 的 n 次多项式.

第四章　线性空间

在第一章和第二章中，我们引进了 n 维向量空间和矩阵的概念. 在搞清这两章的主要内容之后，读者就会发现，线性方程组的理论实际上可以看作是 n 维向量空间和矩阵的一般理论的一个应用. 这说明深入地研究 n 维向量空间和矩阵在理论上是很有意义的. 从本章开始，我们将把 n 维向量空间和矩阵的理论提到中心的地位上来，把它们当作我们的主要研究对象.

在科学上研究某个问题时，一般总是先从具体事物的研究入手，从中概括出某些具有普遍意义的规律，然后把它们抽象化，提升为更一般的理论. 在前面我们研究了 n 维向量空间和矩阵这两种较为具体的事物，得到了许多重要的结论(用命题和定理的形式表达出来). 在这一章和下一章，我们要把这两方面的东西从理论上再加以抽象化，使我们对它们的研究深入一步，从而使它们可以应用到更广泛的领域中去.

§1　线性空间的定义

线性空间的概念是 n 维向量空间概念的抽象和提高. 为了讲清这一重要概念，我们首先回顾一下集合论中的若干术语和记号.

集合论的若干基本概念

在数学上，把某些被考察的对象的全体当作一个整体看待，称为一个**集合**. 例如，我们把数域 K 上的 n 维向量(即其分量均属于数域 K)的全体看成一个集合. 在第一章§2中，我们在这个集合中定义了加法和数乘运算，然后称之为数域 K 上的 n 维向量空

间. 又如，我们把数域 K 上全体 $m\times n$ 矩阵看成一个集合，这个集合通常记为 $M_{m,n}(K)$. 而数域 K 上全体 n 阶方阵也成一集合，这个集合在本书中固定用记号 $M_n(K)$ 表示.

如果 a 是集合 M 的一个成员，我们就说 a 是集合 M 的一个元素，记作 $a\in M$，读作"a 属于 M". 如 a 不属于集合 M，则记作 $a\notin M$，读作"a 不属于 M". 如果 M,N 是两个集合，而 N 中的每个元素都属于 M，则称 N 是 M 的一个**子集合**，记作 $N\subseteq M$，读作"M 包含 N".

设 M,N 是两个集合，把它们的元素合并在一起得到一个新集合，称为 M 和 N 的**并集**，记作 $M\cup N$. 而集合 M 和 N 的公共元素的全体也成一集合，称为 M 和 N 的**交集**，记作 $M\cap N$.

例如，有理数集合是实数集合的一个子集合；有理数集合和实数集合的并是实数集合；有理数集合和实数集合的交是有理数集合. 又如，两条直线 l_1 和 l_2 交于 O 点. 命 M 表 l_1 上全体点所成的集合，N 表 l_2 上全体点所成的集合，则 $M\cup N$ 是 l_1 和 l_2 上点的全体所成的集合；而 $M\cap N$ 是仅由点 O 组成的集合（仅有一个元素）.

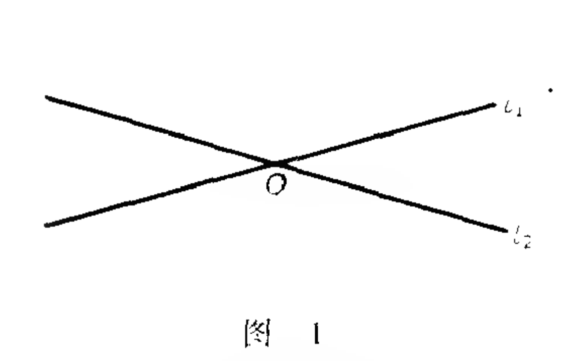


图 1

不包含任何元素的集合称为**空集合**，记作 $\varnothing$. 例如方程 $x^2+1=0$ 的实根所组成的集合是一空集合. 空集合并不是由数 0 组成的集合，而是不包含任何元素的"集合". 按理说，它不应当称为集合，但是使用空集合这一概念对我们讨论许多问题是有好处的. 例如，当集合 M 和 N 没有公共元素时，我们就说它们的交集是空集合，记作 $M\cap N=\varnothing$.

两个集合 M 和 N 相等，是指它们的元素完全一样. 要判断两集合是否相等，通常可以从两个方面进行考察：(i) 是否 $M\subseteq N$，即 M 中每个元素是否都属于 N；(ii) 是否 $N\subseteq M$，即 N 中每个元

素是否都属于 M. 如果上面两条都成立，则可断定 $M=N$.

给定两个集合 M 和 N，如果又给定一个法则 f，使 M 中每一个元素 a 在 f 下都对应于 N 中一个唯一确定的元素，记作 $f(a)$，则称 f 是集合 M 到集合 N 的一个**映射**. $f(a)$ 称为 a 的**象**，而 a 则称为 $f(a)$ 的一个**原象**.

例如，M 是平面上全体点所成的集合，N 是实数的二元有序数组 (x,y) 的全体所成的集合. 取定平面上一个直角坐标系后，就给了一个法则，使平面上每个点，即 M 的每个元素都对应于唯一的一个实数的二元有序数组（即该点的坐标），这就给了 M 到 N 的一个映射. 这个映射是平面解析几何中坐标法的基础. 又如，M 是区间 $(-\infty,+\infty)$ 内全体实数所成的集合，N 是区间 $[-1,1]$ 内全体实数所成的集合，命

$$f(x)=\sin x,\quad x\in(-\infty,+\infty).$$

则 f 是 M 到 N 的一个映射. 不难看出，数学分析中的函数都是从一个集合（函数定义域）到另一个集合（函数值域）的映射.

设 f 是集合 M 到集合 N 的一个映射，下列三种情况是特别重要的：

1. M 中不同元素 a,b 在 f 下的象 $f(a),f(b)$ 也是 N 中的不同元素，此时 f 称为单一映射，简称**单射**；

2. 对 N 中任一元素 b，都至少存在 M 中一个元素 a，使 $f(a)=b$，此时 f 称为**满射**；

3. f 既是单射，又是满射，此时 f 称为**双射**或**一一对应**.

上面举出的第一个例子：平面上的点对应于实数的二元有序数组，是一一对应；而第二个例子（正弦函数）是满射，但不是单射. 因为在区间 $(-\infty,+\infty)$ 内两个不同的实数的正弦函数值有可能相等.

现在设 f 是集合 M 到集合 N 的一个一一对应. 此时任给 N 中一元素 b，它在 M 中恰好存在唯一的一个原象 a. 如果我们定义 N 到 M 的一个映射 g 如下：对 N 中任意元素 b，让它在 g 下对应

于它在 f 下的唯一原象 a，那么容易看出，N 中不同元素在 g 下也对应于 M 中不同元素，即 g 为 N 到 M 的单射，而且 M 中任一元素 a，若 $f(a)=b$，则 b 在 g 下的象即为 a，即 g 为 N 到 M 的满射，因而 g 是 N 到 M 的一一对应. 我们把 g 称为 f 的**逆映射**，记作 f^{-1}. 因此，集合间的双射（一一对应）都存在逆映射.

线性空间的定义

我们首先回想一下，在第一章中，我们把数域 K 上 n 维向量的全体所成的集合，连同它们的加法与数乘运算合称为数域 K 上的 n 维向量空间. 在那里，我们用命题 2.1 概括了它的最基本的性质. 我们当时曾经指出，只要有了该命题所列举的八条性质，那么后面所论述的许多结论就都成立. 具体地说，像向量组的线性相关与线性无关的概念（利用命题 2.2 及其推论中给出的等价定义）以及与之相关的命题 2.3，2.4 及 2.5 等等就都成立. 这一事实启发我们：可以不限于研究 n 元有序数组这种比较具体的东西，而去研究一个更一般的集合. 只要设法在这个集合中确定出加法和数乘（关于某一数域 K）运算，使之满足第一章命题 2.1 所指出的八条规律，那么，前面关于 n 维向量空间的许多结论就可以平行地推移到这一新的研究领域中来了.

例如，考察闭区间 $[a,b]$ 上全体实连续函数 $f(x)$ 所组成的集合，我们记之为 $C[a,b]$. 这个集合内任意两个元素 $f(x)$ 与 $g(x)$，它们作为函数可以有加法运算：$f(x)+g(x)$，而且其和仍为 $[a,b]$ 上的连续函数，也就是仍然属于集合 $C[a,b]$. 对任一实数 k，有乘法 $kf(x)$，且乘以 k 后所得的 $kf(x)\in C[a,b]$. 这样一来，在集合 $C[a,b]$ 的元素之间有加法和与实数的数乘运算. 不难验证，这两种运算具有第一章命题 2.1 所指出的八条基本性质，因而第一章 §2 中的许多基本概念和命题就可以应用到 $C[a,b]$ 上来了. $C[a,b]$ 中的元素和 n 维向量空间中的向量是根本不同的两种东西，但从上面的分析可知，它们之间有着某些共同点. 这些共同点追根究

源是由于它们都具有上面所说的两种运算和八条性质而产生的.

像 $C[a,b]$这样的例子还可以举出很多,从理论上加以概括和抽象化,就得到线性空间的一般性概念.

定义 设 V 是一个非空集合,K 是一个数域. 又设:

(i) 在 V 中定义了一种运算,称为**加法**. 即对 V 中任意两个元素 α 与 β,都按某一法则对应于 V 内唯一确定的一个元素,记之为 $\alpha+\beta$;

(ii) 在 V 中定义了一种运算,称为**数乘**. 即对 V 中任意元素 α 和数域 K 中任意数 k,都按某一法则对应于 V 内唯一确定的一个元素,记之为 $k\alpha$.

且满足下面的八条运算法则:

(1) $\alpha+(\beta+\gamma)=(\alpha+\beta)+\gamma$;

(2) $\alpha+\beta=\beta+\alpha$;

(3) 存在一个元素 $0\in V$,使对一切 $\alpha\in V$,有

$$\alpha+0=\alpha,$$

此元素 0 称为 V 的**零元素**;

(4) 对任一 $\alpha\in V$ 都存在 $\beta\in V$,使

$$\alpha+\beta=0,$$

β 称为 α 的一个**负元素**;

(5) 对数域中的数 1,有 $1\cdot\alpha=\alpha$;

(6) 对任意 $k,l\in K,\alpha\in V$,有

$$(kl)\alpha=k(l\alpha);$$

(7) 对任意 $k,l\in K,\alpha\in V$,有

$$(k+l)\alpha=k\alpha+l\alpha;$$

(8) 对任意 $k\in K,\alpha,\beta\in V$,有

$$k(\alpha+\beta)=k\alpha+k\beta.$$

则称 V 是数域 K 上的一个**线性空间**.

显然,n 维向量空间是线性空间的具体例子. 我们再来举一些例子.

例 1 取集合 V 为三维几何空间中全体向量(有向线段),K 为实数域 $\mathscr{R}$. 则 V 关于向量加法(平行四边形法则)和数乘构成一实数域 $\mathscr{R}$ 上的线性空间.

例 2 取数域 K 上 $m\times n$ 矩阵的全体所成的集合 $M_{m,n}(K)$,加法和数乘为矩阵加法和数乘. 显然,定义中的八条性质是具备的,因此,$M_{m,n}(K)$构成一数域 K 上的线性空间.

例 3 取全体实数所成的集合 $\mathscr{R}$,其加法为普通实数的加法,其与有理数的数乘为普通数的乘法,则 $\mathscr{R}$ 为有理数域 $\mathscr{Q}$ 上的一个线性空间.

例 4 取以数域 K 的元素作系数的一元多项式 $f(X)$的全体所成的集合,记之为 $K[X]$. 加法为通常多项式的加法,与 K 的元素的数乘为通常数与多项式的乘法. 显然,定义中的八条性质是具备的,因而 $K[X]$是数域 K 上的一个线性空间.

例 5 取以数域 K 的元素作系数而次数小于 n 的一元多项式 $f(X)$的全体所成的集合,记之为 $K[X]_n$. 加法与数乘运算同例 4. 显然,$K[X]_n$ 也是数域 K 上的一个线性空间.

如果取 V 为 $K[X]$中全体 n 次多项式(n 为一个固定的正整数)所成的集合,加法与数乘定义同例 4,那么此时 V 不成一线性空间,因为它关于加法运算不封闭(或者更确切地说,这个集合内多项式的加法并不总有意义). 例如:$X^n\in V$, $-X^n\in V$,但$(X^n)+(-X^n)=0\notin V$.

从上面的例子已经可以看出,线性空间的概念比 n 维向量空间的概念具有更大的普遍性,因而,它的应用范围也更广.

因为线性空间是从向量空间抽象出来的,所以我们把线性空间的元素也称为**向量**. 虽然,它已不再具有三维几何空间中向量的几何直观意义,也不再具有 n 维向量空间中向量的 n 元有序数组的具体形式. 线性空间中的零元素称为**零向量**. 一个向量 α 的负元素则称为 α 的**负向量**.

*线性空间的基本关系

设 V 是数域 K 上的一个线性空间. 由于 V 内的加法和数乘运算是以非常抽象的方式给出的,可能跟我们熟悉的数或 n 维向量空间中向量的运算相去甚远,因此,我们已经习以为常的一些东西,严格来讲,都需要重新从逻辑上加以证明. 下面把它们列举出来.

(1) V 中的零向量是唯一的.

设 $0_1, 0_2$ 为 V 的两个零向量,按零向量的定义,有:

$$0_1 = 0_1 + 0_2 = 0_2 + 0_1 = 0_2.$$

(2) V 中任一向量 α 的负向量是唯一的.

设 β, γ 均为 α 的负向量,则

$$\begin{aligned}\beta &= \beta + 0 = \beta + (\alpha + \gamma) = (\beta + \alpha) + \gamma \\ &= 0 + \gamma = \gamma.\end{aligned}$$

α 的唯一负向量今后记作 $-\alpha$,而 $\alpha+(-\beta)$ 记作 $\alpha-\beta$,这就是线性空间中的减法运算.

(3) 消去律. 若 $\alpha+\beta=\alpha+\gamma$,则 $\beta=\gamma$.

以 $-\alpha$ 加等式两端,得

$$-\alpha + (\alpha + \beta) = -\alpha + (\alpha + \gamma).$$

再用结合律以及负向量和零向量的定义,即知 $\beta=\gamma$.

(4) $0 \cdot \alpha=0$; $(-1)\alpha=-\alpha$; $k \cdot 0=0$.

分别证这三个等式.

因为 $0 \cdot \alpha+0 \cdot \alpha=(0+0) \cdot \alpha=0 \cdot \alpha=0 \cdot \alpha+0$,由消去律,即得 $0 \cdot \alpha=0$.

因为 $\alpha+(-1)\alpha=1 \cdot \alpha+(-1)\alpha=(1+(-1))\alpha=0 \cdot \alpha=0$,故 $(-1)\alpha$ 为 α 的负向量,即 $(-1)\alpha=-\alpha$.

因为 $k \cdot 0+k \cdot 0=k(0+0)=k \cdot 0=k \cdot 0+0$,由消去律,即得 $k \cdot 0=0$.

(5) 若 $k\alpha=0$,则 $k=0$ 或 $\alpha=0$.

若 $k\alpha=0$ 且 $k\neq0$,则

$$\alpha = 1 \cdot \alpha = \left(\frac{1}{k}k\right)\alpha = \frac{1}{k}(k\alpha) = \frac{1}{k} \cdot 0 = 0.$$

(6) 若 $k\alpha=\beta, k\neq0$,则 $\alpha=\frac{1}{k}\beta$.

证法同(5).

还有一些更为简单的关系式,这里就不一一写出了.

习 题 一

1. 给定两个集合 M,N,其元素分别为

$$M = \{-1,3,2,1,0\};\ N = \{2,-1,7,0,9,-3\}.$$

试求 $M\cup N$ 和 $M\cap N$.

2. 证明:对任意三个集合 M,N,L,有

$$M \cap (N \cup L) = (M \cap N) \cup (M \cap L);$$
$$M \cup (N \cap L) = (M \cup N) \cap (M \cup L).$$

3. 求 $M_{m,n}(K)\cap M_n(K)$,其中 $m<n$.

4. 求 $M_n(\mathscr{R})\cap M_n(\mathscr{C})$.

5. 命 M 表全体自然数所成的集合;N 表全体正偶数所成的集合. 定义 M 到 N 的映射 f 如下:对任一 $n\in M, f(n)=2n$,证明:f 是 M 到 N 的一一映射.

6. 命 M 表数域 K 上全体二维向量所成的集合;N 表数域 K 上全体三维向量所成的集合. 定义 M 到 N 的映射 f 如下:

$$f:(a_1,a_2)\longmapsto(a_1,a_2,0),$$

证明 f 是 M 到 N 的单一映射,但不是一一映射.

7. 命 M 表数域 K 上全体三维向量所成的集合,定义 M 到 M 自身的映射 f 如下:

$$f:(a_1,a_2,a_3)\longmapsto(a_1,a_2,0),$$

证明 f 既不是单一的,也不是满的.

8. 检验以下集合对于所指定的运算是否构成实数域上的线

性空间：

(1) 闭区间$[a,b]$上全体实连续函数$f(x)$所成的集合$C[a,b]$关于通常函数加法和与实数的乘法；

(2) 设A是n阶实方阵，A的实系数多项式

$$f(A)=a_0A^m+a_1A^{m-1}+\cdots+a_{m-1}A+a_mE\quad(a_i\in\mathscr{R})$$

的全体所成的集合关于矩阵加法和数乘运算；

(3) 全体实对称(反对称、上三角、下三角)矩阵所成的集合关于矩阵的加法和数乘运算；

(4) 平面上不平行于某一非零向量的全部向量所成的集合关于向量的加法和数乘运算；

(5) 全体实数的二元有序数组所成的集合关于下面定义的运算：

$$(a_1,b_1)\oplus(a_2,b_2)=(a_1+a_2,b_1+b_2+a_1a_2),$$

$$k\circ(a,b)=\left(ka,kb+\frac{k(k-1)}{2}a^2\right);$$

(6) 平面上全体向量所成的集合关于通常的向量加法和如下定义的数量乘法：

$$k\circ\alpha=0;$$

(7) 全体正实数所成的集合关于下面定义的运算：

$$a\oplus b=ab,$$

$$k\circ a=a^k.$$

9. 令

$$\omega=\frac{-1+\sqrt{3}\,\mathrm{i}}{2},$$

而$\mathscr{Q}(\omega)$为全体形如

$$a+b\omega\quad(a,b\in\mathscr{Q})$$

的数所成的集合. 定义$\mathscr{Q}(\omega)$内元素的加法为普通数的加法，与有理数k的数乘为普通数的乘法. 证明：$\mathscr{Q}(\omega)$关于上述运算成为有理数域$\mathscr{Q}$上的一个线性空间.

§2 有限维线性空间

设V是数域K上的一个线性空间.前面已经指出:第一章§2中关于n维向量空间的一些基本定义和名词可以平行地推移到V中来.我们这里只简单地把几个最重要的概念重复说一下.

(1) 给定V内一个向量组$\alpha_1,\alpha_2,\cdots,\alpha_s$,如果对$V$内一个向量$\beta$,存在数域$K$内的$s$个数$k_1,k_2,\cdots,k_s$,使

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s,$$

则称β可被向量组$\alpha_1,\alpha_2,\cdots,\alpha_s$**线性表示**.

(2) 给定V内一个向量组$\alpha_1,\alpha_2,\cdots,\alpha_s$,如果存在$K$内不全为零的$s$个数$k_1,k_2,\cdots,k_s$,使

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = 0,$$

则称向量组$\alpha_1,\alpha_2,\cdots,\alpha_s$**线性相关**.如果由

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = 0,$$

必定推出$k_1=k_2=\cdots=k_s=0$,则称向量组$\alpha_1,\alpha_2,\cdots,\alpha_s$**线性无关**.

(3) 给定V内两个向量组

$$\alpha_1,\alpha_2,\cdots,\alpha_r, \qquad (\text{I})$$

$$\beta_1,\beta_2,\cdots,\beta_s, \qquad (\text{II})$$

如果(Ⅰ)中任一向量都能被(Ⅱ)线性表示,反过来,(Ⅱ)中任一向量也都能被(Ⅰ)线性表示,则称两向量组**等价**.

(4) 给定V内一个向量组$\alpha_1,\alpha_2,\cdots,\alpha_s$,如果它有一个部分组$\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$满足如下条件:

(i) 它线性无关;

(ii) 原向量组中任一向量都能被它线性表示.

则称此部分组为原向量组的一个**极大线性无关部分组**.一个向量组的任一极大线性无关部分组中均包含相同数目的向量,其向量数目称为该向量组的**秩**.

在第一章§2证明命题2.3,2.4,2.5时,都仅仅用到该章的

命题 2.1，而没有用到向量的 n 元有序数组的表达形式. 因为命题 2.1 现在已经包含在线性空间的定义中了，所以，对线性空间 V，这些命题也都成立，我们不再重复证明，而将直接加以引用.

但是这里需要强调指出：在 V 中凡涉及数乘运算时，都仅限于 K 中的数. 用 K 以外的数做数乘是没有意义的. 这一点在第一章 § 2 中还没有可能很明确指出来(那时还未给出数域的概念)，现在读者必须注意这一点.

例 1 区间(a,b)内全体可微实函数 $f(x)$所成的集合 $D(a,b)$，其关于通常函数的加法以及与实数的乘法组成实数域上的一个线性空间. 设 λ_1,λ_2 是两个不同的实数，考虑 $D(a,b)$的向量组 $e^{\lambda_1 x},e^{\lambda_2 x}$. 我们来证明这个向量组线性无关. 设

$$k_1 e^{\lambda_1 x} + k e^{\lambda_2 x} = 0,$$

证明 $k_1=k_2=0$. 注意现在做数乘的 k_1,k_2 只能是实的常数. 等式右端的 0 是 $D(a,b)$中的零向量，即常数函数 0. 上面的等式表示在区间(a,b)内左端的函数恒等于零. 把该等式两端对 x 求微商，并与该式联立得方程组

$$\begin{cases} k_1\lambda_1 e^{\lambda_1 x} + k_2\lambda_2 e^{\lambda_2 x} = 0, \\ k_1 e^{\lambda_1 x} + k_2 e^{\lambda_2 x} = 0. \end{cases}$$

以 $x=0$ 代入上面联立方程中，再把 k_1,k_2 看作未知量，解方程组，利用 $\lambda_1\neq\lambda_2$，即知其有唯一解 $k_1=k_2=0$.

线性空间 $D(a,b)$内向量组的线性相关与线性无关概念在微分方程中是很有用的.

例 2 全体形如 $a+b\sqrt{3}$ $(a,b$ 为有理数)的数组成的集合 $\mathscr{Q}(\sqrt{3})$，其关于通常数的加法以及与有理数的乘法组成有理数域 $\mathscr{Q}$ 上的一个线性空间. 考虑 $\mathscr{Q}(\sqrt{3})$内的向量组

$$1, \sqrt{3}, -5.$$

显然，存在不全为零的有理数 k_1,k_2,k_3，使

$$k_1\cdot 1 + k_2\cdot\sqrt{3} + k_3\cdot(-5) = 0$$

(例如取 $k_1=5, k_2=0, k_3=1$). 故这个向量组线性相关. 再考察它的部分组 $1, \sqrt{3}$,那么,对任意有理数 k_1, k_2,由

$$k_1 \cdot 1 + k_2 \cdot \sqrt{3} = 0,$$

必定推出 $k_1=k_2=0$(注意现在做数乘的 k_1, k_2 仅限于有理数,如 k_1, k_2 不全为零,则必定全不为零,于是就有 $\sqrt{3}=-k_1/k_2$ 为有理数,矛盾). 这说明 $1, \sqrt{3}$ 线性无关. 原向量组显然能被它们线性表示,故 $1, \sqrt{3}$ 是向量组 $1, \sqrt{3}, -5$ 的一个极大线性无关部分组. 于是此向量组的秩为 2.

线性空间的维数和基

命题 2.1 设 V 是数域 K 上的一个线性空间. 如果在 V 中存在 n 个线性无关的向量

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n, \qquad (\mathrm{I})$$

使 V 中任一向量均能被(Ⅰ)线性表示,那么,有

(i) 任给 V 内 n 个线性无关向量

$$\eta_1, \eta_2, \cdots, \eta_n, \qquad (\mathrm{II})$$

则 V 中任一向量都能被(Ⅱ)线性表示;

(ii) 如果 $\eta_1, \eta_2, \cdots, \eta_s$ 是 V 的一个线性无关向量组,且 V 内任一向量均能被它们线性表示,则 $s=n$.

证 (i) 任给 $\alpha \in V$,考察向量组

$$\eta_1, \eta_2, \cdots, \eta_n, \alpha.$$

它们能被(Ⅰ)线性表示,其个数$>n$,由第一章命题 2.4,此向量组线性相关. 于是存在 K 内不全为零的数 $k_1, k_2 \cdots, k_n, k$,使

$$k_1\eta_1 + k_2\eta_2 + \cdots + k_n\eta_n + k\alpha = 0.$$

在上式中,如 $k=0$,等式变成

$$k_1\eta_1 + k_2\eta_2 + \cdots + k_n\eta_n = 0.$$

而因为 $\eta_1, \eta_2, \cdots, \eta_n$ 线性无关,必有 $k_1=k_2=\cdots=k_n=0$,与假设矛盾. 故必定 $k\neq 0$,于是

$$\alpha = -\frac{k_1}{k}\eta_1 - \frac{k_2}{k}\eta_2 - \cdots - \frac{k_n}{k}\eta_n.$$

即 α 可被 $\eta_1,\eta_2,\cdots,\eta_n$ 线性表示.

(ii) 根据所给的条件，$\eta_1,\eta_2,\cdots,\eta_s$ 与 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 等价，由第一章命题 2.5 的推论 2，它们的秩相同. 它们又都是线性无关的，其秩分别为 s 与 n，故 $s=n$. ▌

根据命题 2.1，我们可以给出如下一个重要概念：

定义 设 V 是数域 K 上的一个线性空间. 如果 V 中存在 n 个线性无关的向量

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n,$$

使 V 中任一向量均能被此向量组线性表示，则 V 称为 n **维线性空间**，记作 $\dim V=n$. 上述向量组称为 V 的一组**基**. 单由零向量组成的线性空间称为**零空间**. 零空间的维数定义为零.

从命题 2.1 可知：在 n 维线性空间中，任意 n 个线性无关向量都是它的一组基，反之，它的任意一组基也一定恰好含有 n 个向量. 显然，在一个 n 维线性空间中，任意 $n+1$ 个向量都是线性相关的，所以，在一个 n 维线性空间中最多只有 n 个线性无关的向量.

如果在一个线性空间中存在任意多个线性无关的向量，则称之为**无限维线性空间**. 而 n 维线性空间则统称为**有限维线性空间**. 本书主要讨论有限维线性空间.

例 3 考察三维几何空间中全体向量组成的实数域上的线性空间. 从中任取三个不共面的向量 $\varepsilon_1,\varepsilon_2,\varepsilon_3$，由空间解析几何的知识不难知道它们是线性无关的，且任一向量均能被它们线性表示，所以这个线性空间是三维的，而 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 是它的一组基. 这与我们的习惯称呼相一致. 在这个线性空间中，向量组的线性相关与线性无关有直观的几何意义，我们分述如下：

(i) 如果向量 $\alpha_1,\alpha_2,\cdots,\alpha_s(s\geqslant 2)$ 共线，则它们线性相关. 反之，如果两个向量 α_1,α_2 线性相关，则它们是共线的(注意我们这里讨论的向量其起点都设为坐标原点)；

(ii) 如果向量 $\alpha_1,\alpha_2,\cdots,\alpha_s(s\geqslant3)$ 共面，则它们线性相关. 反之，如果三个向量 $\alpha_1,\alpha_2,\alpha_3$ 线性相关，则它们共面；

(iii) 任意向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s(s\geqslant4)$ 必线性相关.

以上结论请读者利用解析几何的知识自己证明.

例 4 考察数域 K 上全体 $m\times n$ 矩阵所组成的线性空间 $M_{m,n}(K)$. 令 $E_{ij}(i=1,2,\cdots,m;j=1,2,\cdots,n)$ 表示第 i 行第 j 列元素为 1，其它元素为零的 $m\times n$ 矩阵. 不难验证向量组 E_{ij} 线性无关，且任一向量(为 $m\times n$ 矩阵)均能被它们线性表示. 故 $E_{ij}(i=1,2,\cdots,m;j=1,2,\cdots,n)$ 是线性空间 $M_{m,n}(K)$ 的一组基，于是 $\dim M_{m,n}(K)=mn$.

例 4 的一个特殊情况是 $m=1$，这就是 n 维向量空间. 它的一组基是

$$\begin{aligned}\varepsilon_1 &= (1,0,\cdots,0),\\ \varepsilon_2 &= (0,1,\cdots,0),\\ &\cdots\cdots\\ \varepsilon_n &= (0,0,\cdots,1).\end{aligned}$$

以后我们用记号 K^n 代表数域 K 上的 n 维向量空间. 特别地，以 $\mathscr{R}^n$ 表实数域上的 n 维向量空间，以 $\mathscr{C}^n$ 表复数域上的 n 维向量空间.

例 5 考察§1 中例 5 的线性空间 $K[X]_n$. 在它里面挑出 n 个向量(即 n 个一元多项式)

$$1,X,X^2,\cdots,X^{n-1},$$

我们证明它们线性无关. 设

$$k_0\cdot1+k_1X+k_2X_2+\cdots+k_{n-1}X^{n-1}=0.$$

和例 1 一样，上面的等式表示 X 用数域 K 内任何数代进去等式都成立. 如果 $k_0,k_1,\cdots,k_{n-1}$ 不全为零，那么等式左端是一个次数 $\leqslant n-1$ 的多项式，它最多只有 $n-1$ 个根. 现在数域 K 内的数都是它的根.而任一数域都包含无穷多个数，这是矛盾的. 因此，必定有 $k_0=k_1=\cdots=k_{n-1}=0$. 另一方面，$K[X]_n$ 内任一向量 $f(X)$ 显然都

能被上述向量组线性表示，于是它们组成 $K[X]_n$ 的一组基. 从而有 $\dim K[X]_n = n$.

例 6 把全体复数所成的集合看作复数域 $\mathscr{C}$ 上的线性空间，则数 1(当作向量看)是它的一组基，故此线性空间是一维的. 如把全体复数所成的集合看作实数域 $\mathscr{R}$ 上的线性空间，则 1,i 是它的一组基，其维数为 2. 这说明线性空间 V 的维数不但与集合 V 有关，而且也与数域 K 有关.

例 7 考察§1 中例 4 的线性空间 $K[X]$. 根据例 5 的分析，不难看出，在向量组

$$1, X, X^2, \cdots, X^n, \cdots$$

中任取有限个都是线性无关的. 所以，$K[X]$是一个无限维线性空间.

向量的坐标

设 V 是数域 K 上的 n 维线性空间，而

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$$

是它的一组基. 任给 $\alpha \in V$，有表示式

$$\alpha = a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n.$$

由于 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 线性无关，根据第一章命题 4.1，这个表达式的系数是唯一确定的. 我们称 $a_1, a_2, \cdots, a_n$ 为 α 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的**坐标**.

为着书写方便，我们借助于矩阵乘法的法则，在形式上写成

$$\alpha = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

当然，这只是一个约定，并非真正的矩阵的乘法. 但我们下面就会看到，矩阵乘法的某些规律，例如结合律，对这种形式的写法也是适用的.

要求一个向量 α 在某一组基下的坐标，只要设法将这个向量表成这组基的线性组合就可以了.

例 8 考虑 n 维线性空间 $K[X]_n$，它的一个向量 $f(X)$ 为 X 的一个一元多项式

$$f(X)=a_0+a_1X+\cdots+a_{n-1}X^{n-1}\quad(a_i\in K).$$

显然，它在基 $1,X,X^2,\cdots,X^{n-1}$ 下的坐标为 $a_0,a_1,a_2,\cdots,a_{n-1}$. 用形式写法，可写为

$$f(X)=(1,X,\cdots,X^{n-1})\begin{pmatrix}a_0\\a_1\\\vdots\\a_{n-1}\end{pmatrix}.$$

容易看出，向量组

$$1,X-a,(X-a)^2,\cdots,(X-a)^{n-1}\quad(a\in K)$$

与向量组 $1,X,X^2,\cdots,X^{n-1}$ 等价，所以它们也组成 $K[X]_n$ 的一组基. 设

$$f(X)=a_0'+a_1'(X-a)+a_2'(X-a)^2+\cdots+a_{n-1}'(X-a)^{n-1}.$$

从数学分析中的泰勒(Taylor)公式可知

$$a_0'=f(a),\ a_k'=\frac{1}{k!}f^{(k)}(a)\quad(k=1,2,\cdots,n-1).$$

这就是 $f(X)$ 在这新的一组基下的坐标. 由此可知，同一个向量在两组不同基下的坐标一般是不相同的.

下面较详细地讨论一下 K^n 中的基与坐标的问题.

(1) 在 K^n 中给定 n 个向量

$$\alpha_1=(a_{11},a_{21},\cdots,a_{n1}),$$
$$\alpha_2=(a_{12},a_{22},\cdots,a_{n2}),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\alpha_n=(a_{1n},a_{2n},\cdots,a_{nn}),$$

要判断它们是否构成 K^n 的一组基，只要判断它们是否线性无关就可以了. 其办法是把这些 n 维向量当作列排成一个矩阵 A，然后

用矩阵消元法求其秩.或利用第三章命题 2.1,从其行列式 $|A|$ 是否为零来判断.

（2）设前面给的向量 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 已经构成 K^n 的一组基,求向量

$$\beta=(b_1,b_2,\cdots,b_n)$$

在这组基下的坐标时,只要解向量方程

$$x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n=\beta.$$

这等价于下列线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1,\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n.\end{cases}$$

或采用矩阵方程,写成

$$AX=B \tag{1}$$

（B 为方程常数项所成的 $n\times1$ 矩阵).

例 9 在 K^4 中给定向量组

$$\begin{aligned}\alpha_1&=(2,1,0,1),\\ \alpha_2&=(1,-3,2,4),\\ \alpha_3&=(-5,0,-1,-7),\\ \alpha_4&=(1,-6,2,6),\end{aligned}$$

因为

$$|A|=\begin{vmatrix}2&1&-5&1\\1&-3&0&-6\\0&2&-1&2\\1&4&-7&6\end{vmatrix}=27\neq0,$$

所以这四个向量线性无关,构成了 K^4 的一组基.又设

$$\beta=(8,9,-5,0).$$

求 β 在这组基下的坐标,只要解方程组

$$\begin{cases} 2x_1 + x_2 - 5x_3 + x_4 = 8, \\ x_1 - 3x_2 \qquad - 6x_4 = 9, \\ \qquad 2x_2 - x_3 + 2x_4 = -5, \\ x_1 + 4x_2 - 7x_3 + 6x_4 = 0. \end{cases}$$

这个方程组必定有唯一解.实际求得其解为

$$x_1 = 3,\ x_2 = -4,\ x_3 = -1,\ x_4 = 1,$$

故

$$\beta = 3\alpha_1 - 4\alpha_2 - \alpha_3 + \alpha_4 = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)\begin{pmatrix} 3 \\ -4 \\ -1 \\ 1 \end{pmatrix}.$$

基变换与坐标变换

一、线性空间中的基变换

设在 n 维线性空间 V 内给定两组基

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n,$$

$$\eta_1, \eta_2, \cdots, \eta_n,$$

每个 η_i 都能被 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 线性表示.设

$$\eta_1 = t_{11}\varepsilon_1 + t_{21}\varepsilon_2 + \cdots + t_{n1}\varepsilon_n,$$

$$\eta_2 = t_{12}\varepsilon_1 + t_{22}\varepsilon_2 + \cdots + t_{n2}\varepsilon_n,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\eta_n = t_{1n}\varepsilon_1 + t_{2n}\varepsilon_2 + \cdots + t_{nn}\varepsilon_n.$$

再度借助矩阵乘法法则,把上面的公式形式地写成

$$(\eta_1, \eta_2, \cdots, \eta_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)\begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{pmatrix}.$$

命

$$T=\begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{pmatrix}.$$

称 T 为从基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 到基 $\eta_1,\eta_2,\cdots,\eta_n$ 的**过渡矩阵**.

命题 2.2 在数域 K 上的 n 维线性空间 V 内给定一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$. T 是 K 上一个 n 阶方阵. 命

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T.$$

则有

(i) 如果 $\eta_1,\eta_2,\cdots,\eta_n$ 是 V 的一组基,则 T 可逆;

(ii) 如果 T 可逆,则 $\eta_1,\eta_2,\cdots,\eta_n$ 是 V 的一组基.

证 设 $T=(t_{ij})$的列向量组为

$$T_1=\begin{pmatrix} t_{11} \\ t_{21} \\ \vdots \\ t_{n1} \end{pmatrix},\quad T_2=\begin{pmatrix} t_{12} \\ t_{22} \\ \vdots \\ t_{n2} \end{pmatrix},\quad \cdots,\quad T_n=\begin{pmatrix} t_{1n} \\ t_{2n} \\ \vdots \\ t_{nn} \end{pmatrix},$$

则有

$$\begin{gathered} \eta_1=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T_1, \\ \eta_2=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \eta_n=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T_n. \end{gathered}$$

以 $x_1,x_2,\cdots,x_n$ 分别乘上面的第 1,2,$\cdots$,n 个等式,然后相加,注意 x_i 可直接乘到 T_i 上(请读者按形式写法的定义自己验证一下),故有

$$\sum_{i=1}^{n}x_i\eta_i=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\{x_1T_1+x_2T_2+\cdots+x_nT_n\}. \qquad (2)$$

下面分别证明命题中的(i)与(ii).

(i) 用反证法. 如果 T 不可逆,则它的列向量组线性相关. 于是有不全为零的数 $k_1,k_2,\cdots,k_n$,使

$$k_1T_1 + k_2T_2 + \cdots + k_nT_n = 0.$$

在(2)式中取 $x_i = k_i (i=1,2,\cdots,n)$，即得

$$k_1\eta_1 + k_2\eta_2 + \cdots + k_n\eta_n = 0,$$

即 $\eta_1,\eta_2,\cdots,\eta_n$ 线性相关. 这与它们组成 V 的一组基的假设矛盾，故 $T_1,T_2,\cdots,T_n$ 线性无关，于是 T 可逆.

(ii) 按命题 2.1，只要证 $\eta_1,\eta_2,\cdots,\eta_n$ 线性无关，那它们就是一组基. 用反证法. 若此向量组线性相关，则有不全为零的一组数 $k_1,k_2,\cdots,k_n$，使 $\sum_{i=1}^{n} k_i\eta_i = 0$. 在(2)式中取 $x_i = k_i (i=1,2,\cdots,n)$，得

$$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\{k_1T + k_2T_2 + \cdots + k_nT_n\} = 0.$$

因为 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是线性无关的，它的一个线性组合为零时，必定其系数全为零，故

$$k_1T_1 + k_2T_2 + \cdots + k_nT_n = 0.$$

这表示 $T_1,T_2,\cdots,T_n$ 线性相关，即 T 不满秩. 这与 T 可逆矛盾. ▌

根据上面的命题，两组基之间的过渡矩阵是可逆矩阵. 反过来，从一组给定的基出发，借助于某一可逆矩阵 T，就可以获得一组新的基.

二、线性空间中的坐标变换

设 V 中一个向量 α 在第一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的坐标为 $x_1, x_2,\cdots,x_n$，即

$$\alpha = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n = (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

又设 α 在第二组基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的坐标为 $y_1,y_2,\cdots,y_n$，即

$$\alpha = y_1\eta_1 + y_2\eta_2 + \cdots + y_n\eta_n = (\eta_1,\eta_2,\cdots,\eta_n)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

因为

$$\eta_i = \sum_{j=1}^{n} t_{ji}\varepsilon_j \quad (i = 1,2,\cdots,n),$$

故

$$\begin{aligned}\alpha &= \sum_{i=1}^{n} y_i\eta_i = \sum_{i=1}^{n} y_i \sum_{j=1}^{n} t_{ji}\varepsilon_j \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} y_i t_{ji}\varepsilon_j = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} t_{ji}y_i\right)\varepsilon_j, \\ &= \sum_{j=1}^{n} x_j\varepsilon_j.\end{aligned}$$

由于 α 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的坐标是唯一的，故有

$$x_j = \sum_{i=1}^{n} t_{ji}y_i \quad (j = 1,2,\cdots,n).$$

写成矩阵形式，就是

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

若令

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

则

$$X = TY.$$

这就是同一个向量 α 在不同两组基下的坐标之间的关系，我们称它为**坐标变换公式**.

现在把上面两方面的分析综合起来叙述如下：当两组基之间的变换公式为

$$(\eta_1,\eta_2,\cdots,\eta_n) = (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T$$

时，相应的坐标变换公式是

$$X = TY,$$

其中

$$\alpha = (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X = (\eta_1,\eta_2,\cdots,\eta_n)Y.$$

如果我们采用上面介绍的形式写法，那么坐标变换公式的推导可以大大简化：

$$\alpha = (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X = (\eta_1,\eta_2,\cdots,\eta_n)Y.$$

以关系式

$$(\eta_1,\eta_2,\cdots,\eta_n) = (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T$$

代入，得

$$\begin{aligned}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X &= [(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T]Y \\ &= (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)(TY).\end{aligned}$$

由于 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是一组基，线性无关，它们的两个线性组合相等时，对应系数相等，故得

$$X = TY.$$

这就是坐标变换公式. 这个推理过程说明，我们所采用的形式写法确实具有矩阵乘法的结合律. 这只要采用和号写法具体加以验证就可以了. 因此，今后我们将直接使用形式写法的结合律、分配律等来推导公式.

例如，由

$$(\eta_1,\eta_2,\cdots,\eta_n) = (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T,\quad |T| \neq 0,$$

两边右乘 T^{-1}，利用结合律，得

$$\begin{aligned}(\eta_1,\eta_2,\cdots,\eta_n)T^{-1} &= [(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T]T^{-1} \\ &= (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)(TT^{-1}) \\ &= (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n).\end{aligned}$$

故有 $$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n) = (\eta_1,\eta_2,\cdots,\eta_n)T^{-1}.$$

即从基 $\eta_1,\eta_2,\cdots,\eta_n$ 到基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 的过渡矩阵是 T^{-1}. 与此相

应，有

$$Y = T^{-1}X.$$

K^n 中的基变换

我们具体探讨一下如何求 K^n 中两组基之间的过渡矩阵.

设第一组基为

$$\varepsilon_1 = (a_{11}, a_{21}, \cdots, a_{n1}),$$
$$\varepsilon_2 = (a_{12}, a_{22}, \cdots, a_{n2}),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\varepsilon_n = (a_{1n}, a_{2n}, \cdots, a_{nn}).$$

第二组基为

$$\eta_1 = (b_{11}, b_{21}, \cdots, b_{n1}),$$
$$\eta_2 = (b_{12}, b_{22}, \cdots, b_{n2}),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\eta_n = (b_{1n}, b_{2n}, \cdots, b_{nn}).$$

而

$$(\eta_1, \eta_2, \cdots, \eta_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)T.$$

按定义，T 的第 i 个列向量的分量是 η_i 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的坐标. 设 T 的第 i 个列向量为 T_i. 以 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 作列向量排成矩阵

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

那么，按照 K^n 中向量坐标的求法(参看上面的(1)式)，T_i 为下列矩阵方程的解

$$AT_i = \begin{pmatrix} b_{1i} \\ b_{2i} \\ \vdots \\ b_{ni} \end{pmatrix} \quad (i = 1, 2, \cdots, n).$$

把这 n 个矩阵等式合并起来，就是

$$AT = B, \tag{3}$$

其中 B 为以 $\eta_1,\eta_2,\cdots,\eta_n$ 作列向量的 n 阶方阵

$$B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}.$$

根据第二章§3中例4的方法，要找满足(3)式的未知矩阵 T，只要写出 $n\times 2n$ 矩阵 (AB)，然后用初等行变换把 A 变成 E，则 B 的位置上出来的就是 T.

例 10 在 K^3 中给定两组基

$$\varepsilon_1 = (1,0,-1),\ \varepsilon_2 = (2,1,1),\quad \varepsilon_3 = (1,1,1);$$

$$\eta_1 = (0,1,1),\quad \eta_2 = (-1,1,0),\ \eta_3 = (1,2,1),$$

求它们之间的过渡矩阵 T.

解 分别以 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 和 η_1,η_2,η_3 作列向量组排成两个矩阵 A 及 B

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix},\quad B = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}.$$

做初等行变换如下：

$$(AB) = \begin{pmatrix} 1 & 2 & 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 2 \\ -1 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 2 & 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 2 \\ 0 & 3 & 2 & 1 & -1 & 2 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 2 & 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & -1 & -2 & -4 & -4 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 2 & 0 & -2 & -5 & -3 \\ 0 & 1 & 0 & -1 & -3 & -2 \\ 0 & 0 & 1 & 2 & 4 & 4 \end{pmatrix}$$

$$\longrightarrow \left(\begin{array}{ccc:ccc} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 & -3 & -2 \\ 0 & 0 & 1 & 2 & 4 & 4 \end{array}\right).$$

于是
$$T = \begin{pmatrix} 0 & 1 & 1 \\ -1 & -3 & -2 \\ 2 & 4 & 4 \end{pmatrix}.$$

当求出了两组基之间的过渡矩阵之后，坐标变换公式也就有了.

*线性空间的同构

设 V_1 与 V_2 是同一数域 K 上的两上线性空间（不限于有限维），如果从 V_1 到 V_2 有一个映射 $f: V_1 \to V_2$，满足如下条件：

(i) f 是一一映射；

(ii) 对任意 $\alpha, \beta \in V_1$，有 $f(\alpha+\beta)=f(\alpha)+f(\beta)$；

(iii) 对任意 $\alpha \in V_1$ 和任意 $k \in K$，有 $f(k\alpha)=kf(\alpha)$.

则称 f 是 V_1 到 V_2 的一个**同构映射**. 如果两个线性空间之间存在一个同构映射，则称这两个线性空间**同构**.

同构的线性空间从代数学的角度看是相同的，因而可以不加区分. 但它们的具体意义可能不同，从其它角度看，它们又是有区别的.

现在设 V 是数域 K 上的 n 维线性空间. 在 V 内取定一组基

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n,$$

则 V 内任一向量 α 可唯一地表示成

$$\alpha = a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n.$$

现在在 V 和 K^n 之间建立一个映射

$$f(\alpha) = (a_1, a_2, \cdots, a_n),$$

即让每个 $\alpha\in V$ 对应于由它在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的坐标 $a_1,a_2,\cdots,a_n$ 确定出的 K^n 中的向量. 由于 α 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的坐标是唯一确定的,所以映射 f 是单一的. 这映射显然也是满的. 因而,f 是 V 到 K^n 的一个一一映射.

另一方面,若

$$\alpha = a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n,$$
$$\beta = b_1\varepsilon_1 + b_2\varepsilon_2 + \cdots + b_n\varepsilon_n,$$

则

$$k\alpha = ka_1\varepsilon_1 + ka_2\varepsilon_2 + \cdots + ka_n\varepsilon_n,$$
$$\alpha + \beta = (a_1 + b_1)\varepsilon_1 + (a_2 + b_2)\varepsilon_2 + \cdots + (a_n + b_n)\varepsilon_n.$$

故

$$\begin{aligned} f(\alpha + \beta) &= (a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n) \\ &= (a_1, a_2, \cdots, a_n) + (b_1, b_2, \cdots, b_n) \\ &= f(\alpha) + f(\beta); \\ f(k\alpha) &= (ka_1, ka_2, \cdots, ka_n) \\ &= k(a_1, a_2, \cdots, a_n) = kf(\alpha). \end{aligned}$$

所以 f 是 V 到 K^n 的同构映射.

由此,每个 n 维线性空间都同构于 n 维向量空间 K^n. 根据这一点即不难证明:数域 K 上任意两个同维数的线性空间都彼此同构(参看下面的习题 15,16,17).

利用同构关系讨论某些问题时,可以使之简化. 例如,证明 V 中两组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 和 $\eta_1,\eta_2,\cdots,\eta_n$ 之间的过渡矩阵 $T=(t_{ij})$ 可逆时,可采用如下证法:

取定一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,则按上面所述,从 V 到 K^n 有一个同构映射 f. 而

$$f(\eta_i) = (t_{1i}, t_{2i}, \cdots, t_{ni}) \quad (i = 1,2,\cdots,n).$$

因为 $\eta_1,\eta_2,\cdots,\eta_n$ 线性无关,不难证明(参看下面习题二第 13 题的(4)),它们在 f 下的象也线性无关. 故 T 的列向量组线性无关,其秩为 n,因而 T 是可逆的.

习 题 二

1. 在习题一的第 8 题(1)中所定义的线性空间 $C[a,b]$中,考察下列向量组是否线性相关,并求它们的秩.

(1) $\cos^2 x, \sin^2 x$;

(2) $\cos^2 x, \sin^2 x, 1$;

(3) $\sin x, \sin\sqrt{2}x$;

(4) $\sin\alpha x, \cos\beta x \quad (\alpha\neq 0, \beta\neq 0)$;

(5) $e^{x}, e^{2x}, \cdots, e^{nx}$.

2. 在线性空间 $K[X]$中,证明向量组

$$1, X, X^2, \cdots, X^n$$

与向量组

$$1, X-a, (X-a)^2, \cdots, (X-a)^n$$

$(a\in K, a\neq 0)$线性等价.

3. 在习题一第 9 题所定义的线性空间 $\mathscr{Q}(\omega)$中判断下列向量组是否线性相关,并求其秩:

(1) $1/2, 3, -7$;

(2) $1, \omega, \omega^2, \omega^3, \omega^4$;

(3) $\dfrac{-1+\sqrt{3}\mathrm{i}}{2}, \dfrac{-1-\sqrt{3}\mathrm{i}}{2}, \sqrt{3}\mathrm{i}$.

4. 求线性空间 $\mathscr{Q}(\omega)$的维数和一组基.

5. 证明线性空间 $D(a,b)$(参看本节例 1)是无限维的.

6. 求下列线性空间的维数和一组基:

(1) 数域 K 上全体对称(反对称、上三角、下三角)矩阵所成的数域 K 上的线性空间;

(2) 数域 K 上全体主对角线元素之和为零的 n 阶方阵关于矩阵加法和数乘所成的 K 上线性空间;

(3) 习题一第 8 题的(7)中的线性空间;

(4) 给定矩阵

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}, \quad \omega = \frac{-1+\sqrt{3}\mathrm{i}}{2},$$

由 A 的实系数多项式 $f(A)$的全体所成的集合关于矩阵加法、数乘所成的实数域上线性空间.

7. 证明下列向量组 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 组成 K^4 的一组基,并求向量 β 在这组基下的坐标:

(1) $\varepsilon_1=(1,1,1,1)$, $\varepsilon_2=(1,1,-1,-1)$,

$\varepsilon_3=(1,-1,1,-1)$, $\varepsilon_4=(1,-1,-1,1)$.

$\beta=(1,2,1,1)$.

(2) $\varepsilon_1=(1,1,0,1)$, $\varepsilon_2=(2,1,3,1)$,

$\varepsilon_3=(1,1,0,0)$, $\varepsilon_4=(0,1,-1,-1)$.

$\beta=(1,2,1,1)$.

8. 给定数域 K 上的一个 n 阶方阵 $A\neq 0$. 设

$$f(\lambda) = a_0\lambda^m + a_1\lambda^{m-1} + \cdots + a_m \quad (a_0 \neq 0, a_i \in K)$$

是使 $f(A)=0$ 的最低次多项式. 设 V 是由系数在 K 内的 A 的多项式的全体关于矩阵加法、数乘所组成的 K 上线性空间,证明:

$$E, A, A^2, \cdots, A^{m-1}$$

是 V 的一组基,从而 $\dim V=m$. 求 V 中向量

$$(A - aE)^k \quad (a \in K, 0 \leqslant k \leqslant m)$$

在这组基下的坐标.

9. 接上题. 证明

$$(A - aE)^k \quad (k = 0,1,2,\cdots,m-1)$$

也是 V 的一组基. 求两组基之间的过渡矩阵 T:

$$(E, A - aE, \cdots, (A - aE)^{m-1}) = (E, A, \cdots, A^{m-1})T.$$

10. 在 K^4 中求由基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 到基 $\eta_1,\eta_2,\eta_3,\eta_4$ 的过渡矩阵,并求向量 β 在所指定的基下的坐标.

(1) $\varepsilon_1=(1,0,0,0)$, $\eta_1=(2,1,-1,1)$,

$\varepsilon_2=(0,1,0,0)$, $\eta_2=(0,3,1,0)$,

$\varepsilon_3=(0,0,1,0)$, $\eta_3=(5,3,2,1)$,

$\varepsilon_4=(0,0,0,1)$. $\eta_4=(6,6,1,3)$.

求 $\beta=(b_1,b_2,b_3,b_4)$在 $\eta_1,\eta_2,\eta_3,\eta_4$ 下的坐标.

(2) $\varepsilon_1=(1,2,-1,0)$, $\eta_1=(2,1,0,1)$,

$\varepsilon_2=(1,-1,1,1)$, $\eta_2=(0,1,2,2)$,

$\varepsilon_3=(-1,2,1,1)$, $\eta_3=(-2,1,1,2)$,

$\varepsilon_4=(-1,-1,0,1)$. $\eta_4=(1,3,1,2)$.

求 $\beta=(1,0,0,0)$在 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 下的坐标.

(3) $\varepsilon_1=(1,1,1,1)$, $\eta_1=(1,1,0,1)$,

$\varepsilon_2=(1,1,-1,-1)$, $\eta_2=(2,1,3,1)$,

$\varepsilon_3=(1,-1,1,-1)$, $\eta_3=(1,1,0,0)$,

$\varepsilon_4=(1,-1,-1,1)$. $\eta_4=(0,1,-1,-1)$.

求 $\beta=(1,0,0,-1)$在 $\eta_1,\eta_2,\eta_3,\eta_4$ 下的坐标.

11. 接上题(1). 求一非零向量 ξ,使它在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 与 $\eta_1,\eta_2,\eta_3,\eta_4$ 下有相同的坐标.

*12. 考察线性空间 $K[X]_n$. 给定 n 个互不相同的数 $a_1,a_2,\cdots,a_n$. 令

$$f(X)=(X-a_1)(X-a_2)\cdots(X-a_n).$$

证明:下列多项式组

$$f_i(X)=\frac{f(X)}{X-a_i}\quad(i=1,2,\cdots,n)$$

组成 $K[X]_n$ 的一组基.

13. 设 V 与 V'是数域 K 上的两个线性空间,f 是 V 到 V'的同构映射. 证明:

(1) $f(0)=0$;

(2) $f(-\alpha)=-f(\alpha)$;

(3) 若 $k_1f(\alpha_1)+k_2f(\alpha_2)+\cdots+k_sf(\alpha_s)=0$,则

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=0;$$

(4) 若 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关(线性相关),则 $f(\alpha_1),f(\alpha_2)$,

$\cdots, f(\alpha_s)$也线性无关(线性相关).

14. 证明:实数域$\mathscr{R}$作为它自身上的线性空间与习题一第8题(7)中的线性空间同构.

15. 证明:线性空间的同构是等价关系,即

(1) 每个线性空间V和自己同构;

(2) 设V_1, V_2是数域K上的两个线性空间,如果存在V_1到V_2的同构映射f,则必存在V_2到V_1的同构映射;

(3) 设V_1, V_2, V_3是数域K上的三个线性空间,如果存在V_1到V_2的同构映射f,从V_2到V_3的同构映射g,则必存在V_1到V_3的同构映射.

16. 设V是数域K上的一个n维线性空间,试构造从K^n到V的同构映射.

17. 证明数域K上同维数的任意两个线性空间都互相同构.

§3 子 空 间

前面两节中我们介绍了线性空间的一些最基本的概念,在这一节里,我们来讨论一些稍为深入一点的问题.

子空间的定义

首先来看一看三维几何空间.我们取定一个直角坐标系$OXYZ$.以O为起点的全体向量关于向量加法和数乘组成实数域$\mathscr{R}$上的一个三维线性空间.考察通过原点O的一张平面M上的全体向量所成的集合,我们就用M来表示这个集合.M内两个向量相加仍在M内,M内一个向量乘以实数k后仍在M内.加法和数乘自然满足线性空间定义中的八条性质.所以,M也是实数域上的一个线性空间.其加法和数乘运算是沿用的三维几何空间中向量的加法和数乘运算,这个线性空间是二维的(M内任意不共线的两个向量都是它的一组基).同样地,考察过原点O的任一直

线 l 上全体向量所成的集合，记作 L. L 关于三维几何空间中向量的加法和数乘也组成实数域上的一个线性空间，其维数是 1.

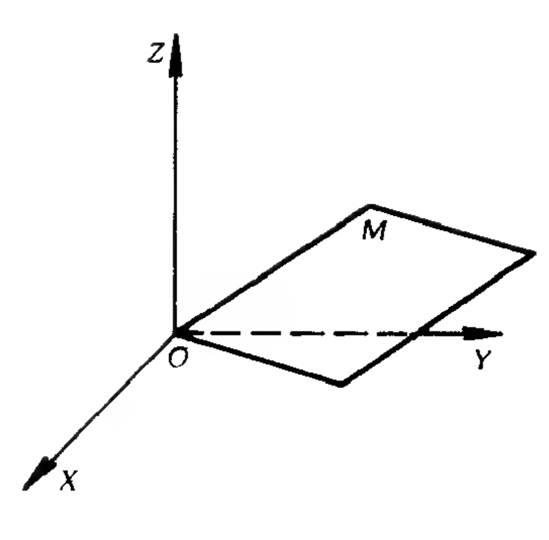


图 2

这些讨论启发我们去考察一般线性空间中的与此相类似的现象.

定义 设 V 是数域 K 上的一个线性空间，M 是 V 的一个非空子集. 如果 M 关于 V 内的加法与数乘运算也组成数域 K 上的一个线性空间，则 M 称为 V 的一个**子空间**.

命题 3.1 线性空间 V 的一个非空子集 M 是一个子空间的充分必要条件是，它满足以下两个条件：

(i) 它对加法封闭，即对 M 内任意两个向量 α,β，有 $\alpha+\beta\in M$；

(ii) 它对数乘运算封闭，即对任一 $\alpha\in M$ 和任一 $k\in K$，有 $k\alpha\in M$.

证 条件的必要性是显然的，我们证明充分性. 设条件(i)和(ii)满足. 我们只要证明：零向量属于 M；M 中任一向量 α 的负向量 $-\alpha$ 属于 M. 那么，线性空间定义中的八条性质就都具备，于是 M 就是子空间了. 因为 M 非空，必有某个 $\alpha\in M$. 由条件(ii)知，$0\cdot\alpha=0\in M$. 另一方面，对任意 $\alpha\in M$，$(-1)\cdot\alpha=-\alpha\in M$. 至此命题得证. ▌

现在设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是 V 内一个向量组，作它们的所有可能的线性组合，由此所得的 V 的子集记作 $L(\alpha_1,\alpha_2,\cdots,\alpha_s)$. 如果采用集合论中惯用的记号，它可写作

$$L(\alpha_1,\alpha_2\cdots,\alpha_s)$$

$$= \{k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s \mid k_i \in K, i = 1,2,\cdots,s\}.$$

这个记号的意思是：集合 $L(\alpha_1,\alpha_2,\cdots,\alpha_s)$的元素是由花括号中所表达出来的向量组成的，而其中线性组合的各个系数 $k_i(i=1,2,\cdots,s)$各自遍历数域 K 中的一切元素. 显然，子集 $L(\alpha_1,\alpha_2,\cdots,\alpha_s)$满足命题 3.1 中的条件(i)和(ii)，故它是 V 的一个子空间，这个子空间称为**由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 生成的子空间**.

下面我们举几个子空间的具体例子.

例 1 $K[X]_n$ 是 $K[X]$的一个子空间. 因为 $K[X]_n$ 作为 $K[X]$的一个非空子集，显然满足命题 3.1 中的条件(i)与(ii).

例 2 给定数域 K 上的一个齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0, \end{cases} \quad (a_{ij} \in K).$$

它的全体解向量是 K^n 中的一个非空子集 M. 根据第一章 §4 中指出的齐次线性方程组解的两条性质，M 满足命题 3.1 中的条件(i)与(ii)，故 M 是 K^n 的一个子空间. 这个子空间称为该齐次线性方程组的**解空间**. 此齐次线性方程组的任一组基础解系就是线性空间 M 的一组基. 如果方程组系数矩阵的秩为 r，则

$$\dim M = n - r.$$

例 3 考察数域 K 上全体 n 阶方阵所成的线性空间 $M_n(K)$. 设 M 是 $M_n(K)$内主对角线元素之和为零的方阵所成的子集，采用集合论的记号，写作

$$M = \{(a_{ij}) \in M_n(K) \mid a_{11} + a_{22} + \cdots + a_{nn} = 0\}.$$

显然，M 也满足命题 3.1 中的条件(i)与(ii)，故 M 是 $M_n(K)$的一个子空间.

应当指出，对于一个线性空间 V，单由它的零向量组成的子集也满足命题 3.1 中的条件，所以它也是 V 的一个子空间，称为**零子空间**，记作$\{0\}$. 另外，V 本身也认为是自身的一个子空间. 这两

个子空间称为 V 的**平凡子空间**. 除此之外的子空间称为**非平凡子空间**.

命题 3.2 设 V 是数域 K 上的 n 维线性空间, M 是 V 的一个非零子空间, 则 M 内任一组基都可以扩充成 V 的一组基.

证 根据命题 2.1, n 维线性空间中至多有 n 个线性无关的向量, 于是 M 内线性无关向量的个数也不能大于 n, 故 M 也是有限维线性空间. 设 $\dim M=r$, 且 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r$ 为 M 的一组基. 于是 $M=L(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r)$.

若 $M=V$, 则 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r$ 已经是 V 的一组基. 如 $M\neq V$, 则存在 $\varepsilon_{r+1}\in V$, 而 $\varepsilon_{r+1}\overline{\in} M$. 我们证明 $\varepsilon_1,\varepsilon_2,,\cdots,\varepsilon_r,\varepsilon_{r+1}$ 必线性无关. 因若

$$k_1\varepsilon_1+\cdots+k_r\varepsilon_r+k_{r+1}\varepsilon_{r+1}=0,$$

首先必有 $k_{r+1}=0$(否则 ε_{r+1} 可被 $\varepsilon_1,\cdots,\varepsilon_r$ 线性表示, 于是 $\varepsilon_{r+1}\in L(\varepsilon_1,\cdots,\varepsilon_r)=M$, 与假设矛盾). 故

$$k_1\varepsilon_1+\cdots+k_r\varepsilon_r=0.$$

再由 $\varepsilon_1,\cdots,\varepsilon_r$ 线性无关推得 $k_1=\cdots=k_r=0$.

如果 $L(\varepsilon_1,\cdots,\varepsilon_r,\varepsilon_{r+1})\neq V$, 再继续上述步骤. 经过 $n-r$ 次这种步骤后, 得到 V 内 n 个线性无关的向量 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$, 由命题 2.1 知它们组成 V 的一组基. ▌

推论 一个有限维线性空间 V 内任意 r 个线性无关的向量 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 都可以扩充为 V 的一组基.

证 在命题 3.2 中取 $M=L(\alpha_1,\alpha_2,\cdots,\alpha_r)$ 即可. ▌

子空间的交与和

定义 设 M_1,M_2 为线性空间 V 的两个子空间, 称 $M_1\cap M_2$ 为它们的**交**. 又命

$$M=\{\alpha_1+\alpha_2\mid\alpha_1\in M_1,\alpha_2\in M_2\},$$

称为 M_1 与 M_2 的**和**, 记作 M_1+M_2.

两个子空间的交是它们的公共部分. 两个子空间的和是分别

由两个子空间中各任取一个向量相加所组成的集合. 因为 M_1 和 M_2 中都包含零向量,所以显然有

$$M_1 \subseteq M_1 + M_2;\quad M_2 \subseteq M_1 + M_2.$$

但要注意 M_1+M_2 和 $M_1 \cup M_2$ 不同. 后者只是把两个子空间的向量简单地聚拢在一起成为一个新的集合而已,它们的向量之间并不相加.

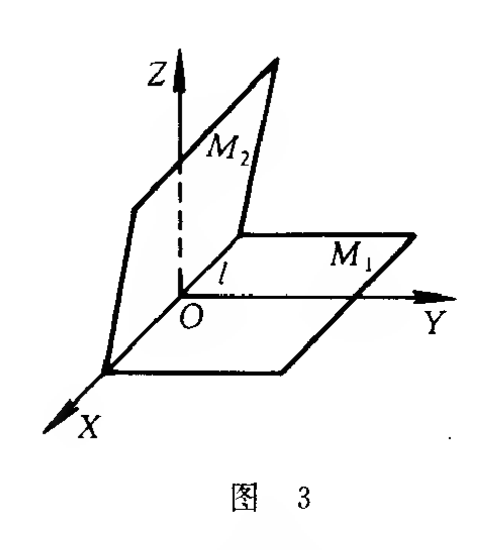


图 3

例 4 考察三维几何空间中过坐标原点 O 的两张不重合的平面(参看图 3). 它们上面的向量分别组成两个子空间 M_1, M_2. 不难看出,M_1+M_2 是整个三维几何空间. 而 $M_1 \cap M_2$ 则是两平面的交线 l 上全体向量组成的一维子空间.

例 5 考察三维几何空间中过坐标原点 O 的两条不重合的直线 l_1 和 l_2. 它们上面的向量分别组成两个一维子空间 L_1 和 L_2. 不难看出,L_1+L_2 是过 l_1 和 l_2 的一张平面上的全体向量组成的二维子空间. 而 $L_1 \cap L_2$ 只有零向量,为零子空间$\{0\}$(参看图 4).

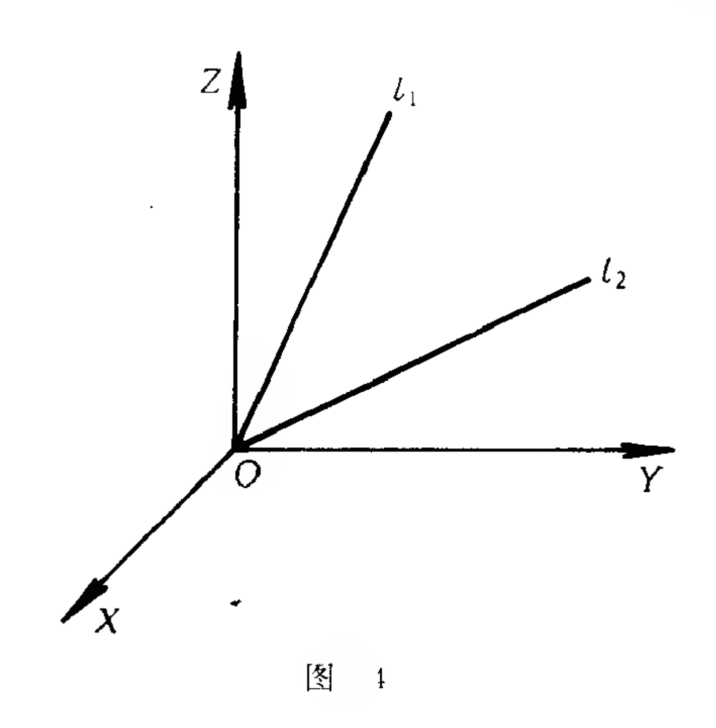


图 4

例 6 给定数域 K 上的两个齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0; \end{cases} \tag{1}$$

$$\begin{cases} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n = 0, \\ b_{21}x_1 + b_{22}x_2 + \cdots + b_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_{l1}x_1 + b_{l2}x_2 + \cdots + b_{ln}x_n = 0; \end{cases} \tag{2}$$

方程组(1)的解空间设为 M_1,方程组(2)的解空间设为 M_2,它们都是 K^n 的子空间. $M_1 \cap M_2$ 中每个向量既是方程组(1)的解向量,又是方程组(2)的解向量. 所以 $M_1 \cap M_2$ 就是把方程组(1)与(2)并在一起得到一个 n 个未知量,$m+l$ 个方程的新方程组的解空间.

在上面的例子中,M_1+M_2 与 $M_1 \cap M_2$ 都恰好还是 V 的子空间. 这一点并不是偶然的,而是一个普遍规律.

命题 3.3 设 M_1, M_2 为线性空间 V 的两个子空间,则它们的交 $M_1 \cap M_2$ 与和 M_1+M_2 仍是 V 的子空间.

证 (i) 证 $M_1 \cap M_2$ 是子空间. 首先,$0 \in M_1 \cap M_2$,故 $M_1 \cap M_2$ 非空. 设 $\alpha, \beta \in M_1 \cap M_2$,则因 $\alpha, \beta \in M_1$,故 $\alpha+\beta \in M_1$;同理,$\alpha+\beta \in M_2$. 于是 $\alpha+\beta \in M_1 \cap M_2$. 又,对任一 $\alpha \in M_1 \cap M_2$ 及任一 $k \in K$,因 $\alpha \in M_1$,有 $k\alpha \in M_1$;同理,$k\alpha \in M_2$. 于是 $k\alpha \in M_1 \cap M_2$. 根据命题 3.1,$M_1 \cap M_2$ 为 V 的子空间.

(ii) 证 M_1+M_2 为子空间. 首先,$0 \in M_1+M_2$,故 M_1+M_2 非空. 若 $\alpha, \beta \in M_1+M_2$,则有

$$\alpha = \alpha_1 + \alpha_2 \quad (\alpha_1 \in M_1, \alpha_2 \in M_2);$$
$$\beta = \beta_1 + \beta_2 \quad (\beta_1 \in M_1, \beta_2 \in M_2).$$

于是,有

$$\alpha + \beta = (\alpha_1 + \beta_1) + (\alpha_2 + \beta_2),$$

其中 $\alpha_1+\beta_1 \in M_1$,$\alpha_2+\beta_2 \in M_2$,故 $\alpha+\beta \in M_1+M_2$. 又,对任一 $\alpha \in$

M_1+M_2 和任一 $k\in K$，因

$$\alpha=\alpha_1+\alpha_2\quad(\alpha_1\in M_1,\alpha_2\in M_2),$$

故
$$k\alpha=k\alpha_1+k\alpha_2,$$

其中 $k\alpha_1\in M_1$，$k\alpha_2\in M_2$，即 $k\alpha\in M_1+M_2$. 再由命题 3.1 知 M_1+M_2 亦为 V 的子空间. ▌

有了两个子空间的交与和后，可以类似地定义多个子空间的交与和. 设 $M_1,M_2,\cdots,M_k$ 为 V 的 k 个子空间，它们的交定义为

$$M_1\cap M_2\cap\cdots\cap M_k,$$

即这 k 个子空间的公共向量所组成的集合. 显然，它是 V 的一个子空间.

又定义

$$M=\{\alpha_1+\alpha_2\cdots+\alpha_k\,|\,\alpha_i\in M_i,i=1,2,\cdots,k\},$$

称为这 k 个子空间的和，记为 $M_1+M_2+\cdots+M_k$. 显然，它也是一个子空间.

现在我们证明一个重要的结果.

定理 1 设 M_1,M_2 是线性空间 V 的两个有限维子空间，则有

$$\dim(M_1+M_2)=\dim M_1+\dim M_2-\dim M_1\cap M_2.$$

证 $M_1\cap M_2$ 既是 M_1 的子空间，又是 M_2 的子空间. 在 $M_1\cap M_2$ 内取定一组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r.$$

按命题 3.2，它可以扩充成 M_1 的一组基

$$\varepsilon_1,\cdots,\varepsilon_r,\alpha_1,\cdots,\alpha_s,$$

也可以扩充成 M_2 的一组基

$$\varepsilon_1,\cdots,\varepsilon_r,\beta_1,\cdots,\beta_t$$

（当 $M_1\cap M_2=\{0\}$ 时，取 $r=0$ 即可）. 我们证明向量组

$$\varepsilon_1,\cdots,\varepsilon_r,\alpha_1,\cdots,\alpha_s,\beta_1,\cdots,\beta_t\qquad\text{(I)}$$

是 M_1+M_2 的一组基.

(i) 对任一 $\alpha\in M_1+M_2$，有

$$\alpha=\alpha_1+\alpha_2\quad(\alpha_1\in M_1,\alpha_2\in M_2).$$

而

$$\alpha_1 = k_1\varepsilon_1 + \cdots + k_r\varepsilon_r + a_1\alpha_2 + \cdots + a_s\alpha_2,$$
$$\alpha_2 = l_1\varepsilon_1 + \cdots + l_r\varepsilon_r + b_1\beta_1 + \cdots + b_t\beta_t.$$

故有

$$\alpha = \alpha_1 + \alpha_2 = (k_1 + l_1)\varepsilon_1 + \cdots + (k_r + l_r)\varepsilon_r + a_1\alpha_1 + \cdots + a_s\alpha_s + b_1\beta_1 + \cdots + b_t\beta_t.$$

即 α 可被向量组(I)线性表示.

(ii) 再证向量组(I)线性无关. 若

$$k_1\varepsilon_1 + \cdots + k_r\varepsilon_r + a_1\alpha_1 + \cdots + a_s\alpha_s + b_1\beta_1 + \cdots + b_t\beta_t = 0, \tag{3}$$

移项,得

$$k_1\varepsilon_1 + \cdots + k_r\varepsilon_r + a_1\alpha_1 + \cdots + a_s\alpha_s = - b_1\beta_1 - \cdots - b_t\beta_t.$$

因为

$$k_1\varepsilon_1 + \cdots + k_r\varepsilon_r + a_1\alpha_1 + \cdots + a_s\alpha_s \in M_1,$$
$$- b_1\beta_1 - \cdots - b_t\beta_t \in M_2,$$

而两者相等,故

$$\gamma = k_1\varepsilon_1 + \cdots + k_r\varepsilon_r + a_1\alpha_1 + \cdots + a_s\alpha_s \in M_1 \cap M_2.$$

于是 γ 可被 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r$ 线性表示,即

$$\gamma = f_1\varepsilon_1 + \cdots + f_r\varepsilon_r.$$

又已知 $$\gamma = -b_1\beta_1 - \cdots - b_t\beta_t,$$

两式相减,得

$$0 = f_1\varepsilon_1 + \cdots + f_r\varepsilon_r + b_1\beta_1 + \cdots + \beta_t\beta_t.$$

由于 $\varepsilon_1,\cdots,\varepsilon_r,\beta_1,\cdots,\beta_t$ 线性无关,故 $b_1 = \cdots = b_t = 0$. 代入(3)式,得

$$k_1\varepsilon_1 + \cdots + k_r\varepsilon_r + a_1\alpha_1 + \cdots + a_s\alpha_s = 0.$$

再由 $\varepsilon_1,\cdots,\varepsilon_r,\alpha_1,\cdots,\alpha_s$ 线性无关,推得 $k_1 = \cdots = k_r = a_1 = \cdots = a_s = 0$. 故向量组(I)线性无关.

综合(i)与(ii)知向量组(I)是 M_1+M_2 的一组基，因此，$\dim(M_1+M_2)=r+s+t$，而 $\dim(M_1\cap M_2)=r$，$\dim M_1=r+s$，$\dim M_2=r+t$，于是有

$$\dim(M_1+M_2)=\dim M_1+\dim M_2-\dim M_1\cap M_2. \blacksquare$$

这个定理中的公式称为**维数公式**.

例 7 在 K^3 中给定两个向量组

$$\alpha_1=(-1,1,0),\ \alpha_2=(1,1,1);$$

$$\beta_1=(-1,3,0),\ \beta_2=(-1,1,-1),$$

求 $L(\alpha_1,\alpha_2)+L(\beta_1,\beta_2)$ 和 $L(\alpha_1,\alpha_2)\cap L(\beta_1,\beta_2)$ 的维数和一组基.

解 (i) 求 $L(\alpha_1,\alpha_2)+L(\beta_1,\beta_2)$ 的维数和一组基. 因为

$$L(\alpha_1,\alpha_2)+L(\beta_1,\beta_2)=L(\alpha_1,\alpha_2,\beta_1,\beta_2),$$

只要求 $\alpha_1,\alpha_2,\beta_1,\beta_2$ 的一个极大线性无关部分组就可以了. 应用第一章§3所讲的办法

$$\begin{bmatrix} -1 & 1 & 0 & \alpha_1 \\ 1 & 1 & 1 & \alpha_2 \\ -1 & 3 & 0 & \beta_1 \\ -1 & 1 & -1 & \beta_2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 1 & \alpha_2 \\ -1 & 1 & 0 & \alpha_1 \\ -1 & 3 & 0 & \beta_1 \\ -1 & 1 & -1 & \beta_2 \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} 1 & 1 & 1 & \alpha_2 \\ 0 & 2 & 1 & \alpha_1+\alpha_2 \\ 0 & 4 & 1 & \beta_1+\alpha_2 \\ 0 & 2 & 0 & \beta_2+\alpha_2 \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} 1 & 1 & 1 & \alpha_2 \\ 0 & 2 & 1 & \alpha_1+\alpha_2 \\ 0 & 0 & 1 & 2\alpha_1+\alpha_2-\beta_1 \\ 0 & 0 & 0 & \alpha_1+\alpha_2-\beta_1+\beta_2 \end{bmatrix}.$$

所以向量组 $\alpha_1,\alpha_2,\beta_1,\beta_2$ 的秩为 3. 又因为

$$\alpha_1+\alpha_2-\beta_1+\beta_2=0, \tag{4}$$

即 $\beta_1=\alpha_1+\alpha_2+\beta_2$，故 $\alpha_1,\alpha_2,\beta_2$ 为 $L(\alpha_1,\alpha_2)+L(\beta_1,\beta_2)$ 的一组基，

其维数为3.

(ii) 求 $L(\alpha_1,\alpha_2)\cap L(\beta_1,\beta_2)$ 的维数和一组基. 因为

$$\dim L(\alpha_1,\alpha_2) = \dim L(\beta_1,\beta_2) = 2,$$

从(i)的结果,利用维数公式知

$$\dim L(\alpha_1,\alpha_2) \cap L(\beta_1,\beta_2) = 1.$$

又从(4)式得

$$\alpha_1 + \alpha_2 = \beta_1 - \beta_2 \in L(\alpha_1,\alpha_2) \cap L(\beta_1,\beta_2).$$

而

$$\alpha_1+\alpha_2=(0,2,1)\neq 0,$$

故 $\alpha_1+\alpha_2$ 为 $L(\alpha_1,\alpha_2)\cap L(\beta_1,\beta_2)$ 的一组基.

习 题 三

1. 考虑线性空间 $M_n(K)$. 设 $A\in M_n(K)$,证明 A 的系数属于 K 的多项式 $f(A)$ 的全体所成的集合 M 是 $M_n(K)$ 的一个子空间.

2. 设 A 是数域 K 上的一个 n 阶方阵,证明:存在一个系数属于 K 的非零多项式 $f(\lambda)$,使 $f(A)=0$.

3. 命 V 表平面上起点在坐标原点 O 的全体向量所成的实数域上二维线性空间,M 表 V 中坐标(在某一直角坐标系下)为有理数的向量的全体所成的子集,问 M 是否为 V 的子空间?

4. 设 $A\in M_n(K)$.

(1) 证明:与 A 可交换的 n 阶方阵的全体组成 $M_n(K)$ 的一个子空间. 记此子空间为 $C(A)$;

(2) 给定对角矩阵

$$A = \begin{pmatrix} 1 & & & \\ & 2 & & \\ & & \ddots & \\ & & & n \end{pmatrix},$$

求 $C(A)$ 的维数和一组基.

5. 接上题. 取 $n=3$,令

$$A=\begin{pmatrix}1&0&0\\0&1&0\\3&1&2\end{pmatrix},$$

求 $C(A)$的维数和一组基.

6. 在实数域上线性空间 $C[-\pi,\pi]$(参看习题一第 8 题的(1))中由向量组

$$1,\ \cos x,\ \cos 2x,\ \cos 3x$$

生成一个子空间 $L(1,\cos x,\cos 2x,\cos 3x)$,求此空间的维数.

7. 证明:有限维线性空间 V 的任一子空间 M 都可以看作是 V 内一个有限向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 生成的子空间.

8. 如果 $k_1\alpha+k_2\beta+k_3\gamma=0$,且 $k_1k_3\neq 0$,证明:

$$L(\alpha,\beta)=L(\beta,\gamma).$$

9. 设 V_1,V_2 都是线性空间 V 的子空间,且 $V_1\subseteq V_2$. 证明:若 $\dim V_1=\dim V_2$,则 $V_1=V_2$.

10. 在 K^4 中求由向量 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 生成的子空间的基与维数.

(1) $\alpha_1=(2,1,3,1)$, $\alpha_2=(1,2,0,1)$,
$\alpha_3=(-1,1,-3,0)$, $\alpha_4=(1,1,1,1)$.

(2) $\alpha_1=(2,1,3,-1)$, $\alpha_2=(-1,1,-3,1)$,
$\alpha_3=(4,5,3,-1)$, $\alpha_4=(1,5,-3,1)$.

11. 在 K^4 中求齐次线性方程组

$$\begin{cases}3x_1+2x_2-5x_3+4x_4=0,\\3x_1-x_2+3x_3-3x_4=0,\\3x_1+5x_2-13x_3+11x_4=0\end{cases}$$

的解空间的基与维数.

12. 求由下列向量 α_i 所生成的子空间与由下列向量 β_i 生成的子空间的交与和的维数和一组基.

(1) $\alpha_1=(1,2,1,0)$, $\beta_1=(2,-1,0,1)$,
$\alpha_2=(-1,1,1,1)$. $\beta_2=(1,-1,3,7)$.

(2) $\alpha_1=(1,1,0,0)$, $\beta_1=(0,0,1,1)$,

$\alpha_2=(1,0,1,1)$. $\beta_2=(0,1,1,0)$.

(3) $\alpha_1=(1,2,-1,-2)$, $\beta_1=(2,5,-6,-5)$,

$\alpha_2=(3,1,1,1)$, $\beta_2=(-1,2,-7,3)$.

$\alpha_3=(-1,0,1,-1)$.

13. 考察三维几何空间.

(1) 问所有终点在同一个平面上的向量是否构成一个子空间?

(2) 设有过原点的三条直线,这三条直线上的全部向量分别组成三个子空间 L_1,L_2,L_3. 问:$L_1+L_2,L_1+L_2+L_3$ 能构成哪些类型的子空间? 试全部列举出来.

14. 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 n 维线性空间 V 的一组基,A 是一个 $n\times s$矩阵. 令

$$(\beta_1,\beta_2,\cdots,\beta_s)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A.$$

证明:$L(\beta_1,\beta_2,\cdots,\beta_s)$的维数等于 A 的秩 $r(A)$.

15. 命 N 表齐次线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0,\\a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0,\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0\end{cases}$$

的解空间,命 M_i 表齐次线性方程

$$a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n=0\quad(i=1,2,\cdots,m)$$

的解空间,证明:$N=M_1\cap M_2\cap\cdots\cap M_m$.

* §4 子空间的直和、商空间

在上一节中,我们介绍了两个以及多个子空间的和的概念. 在这一节里,我们进一步探讨其中最为重要的一种类型:直和.

子空间的直和

我们先考察三维几何空间 V. 设过坐标原点 O 的两张不重合的平面分别代表 V 的两个子空间 M_1 与 M_2(参看图 5). 显然, M_1

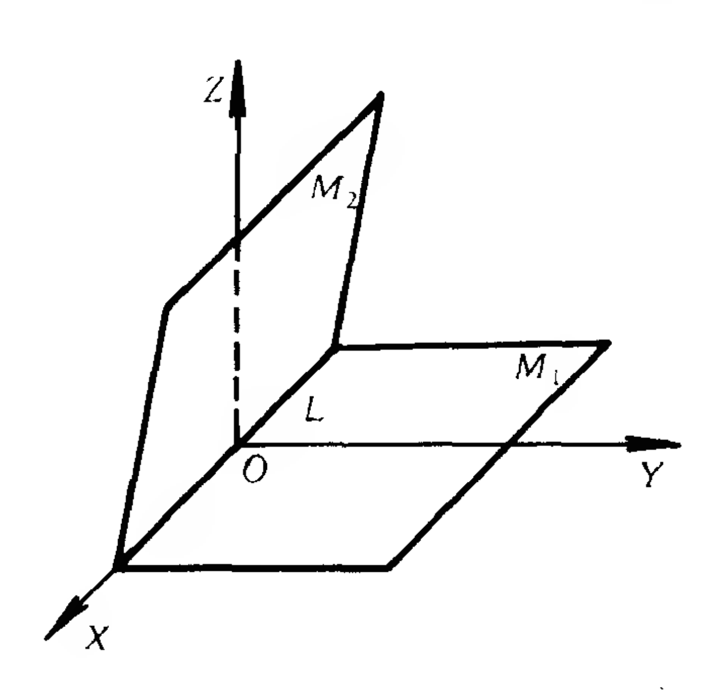


图 5

$+M_2=V$. 因此, V 内任一向量 α 可表为

$$\alpha = \alpha_1 + \alpha_2 \quad (\alpha_1 \in M_1, \alpha_2 \in M_2).$$

但现在这个表示方法不是唯一的. 就是说, 我们可以找到 α 的另一表示方法

$$\alpha = \alpha_1' + \alpha_2' \quad (\alpha_1' \in M_1, \alpha_2' \in M_2),$$

且 $\alpha_1' \neq \alpha_1, \alpha_2' \neq \alpha_2$. 最简单的例子是考察 V 中的零向量. 设 β 为 $M_1 \cap M_2 = L$ 中任一向量. 因为 $L \neq \{0\}$, 可令 $\beta \neq 0$, 而 $\beta \in M_1, -\beta \in M_2$, 我们有

$$0 = 0 + 0 = \beta + (-\beta).$$

即 V 中零向量至少有两种(实际上有无穷多种)方式表为 M_1 和 M_2 中向量之和的形式.

只要零向量表法不唯一, 那么, 任意向量的表法也不唯一. 因为: 设 α 有一个表示

$$\alpha = \alpha_1 + \alpha_2 \quad (\alpha_1 \in M_1, \alpha_2 \in M_2).$$

那么，有

$$\begin{aligned}\alpha &= \alpha + 0 = (\alpha_1 + \alpha_2) + (\beta + (-\beta)) \\ &= (\alpha_1 + \beta) + (\alpha_2 - \beta),\end{aligned}$$

其中 $\alpha_1+\beta\in M_1$，$\alpha_2-\beta\in M_2$，且 $\alpha_1+\beta\neq\alpha_1$，$\alpha_2-\beta\neq\alpha_2$.

我们再考察另一种情况. 设 V 中过坐标原点 O 的一张平面上的全体向量所成的集合 M，和过 O 点但不在 M 内的直线 l 上的全体向量所成的集合 L 分别代表 V 的两个子空间(参看图 6). 此时也有

$$V = M + L.$$

于是，V 内任一向量 α 可表为

$$\alpha = \alpha_1 + \alpha_2 \quad (\alpha_1 \in M, \alpha_2 \in L),$$

此时表法必定是唯一的. 因若还有另一表法

$$\alpha = \alpha_1' + \alpha_2' \quad (\alpha_1' \in M, \alpha_2' \in L),$$

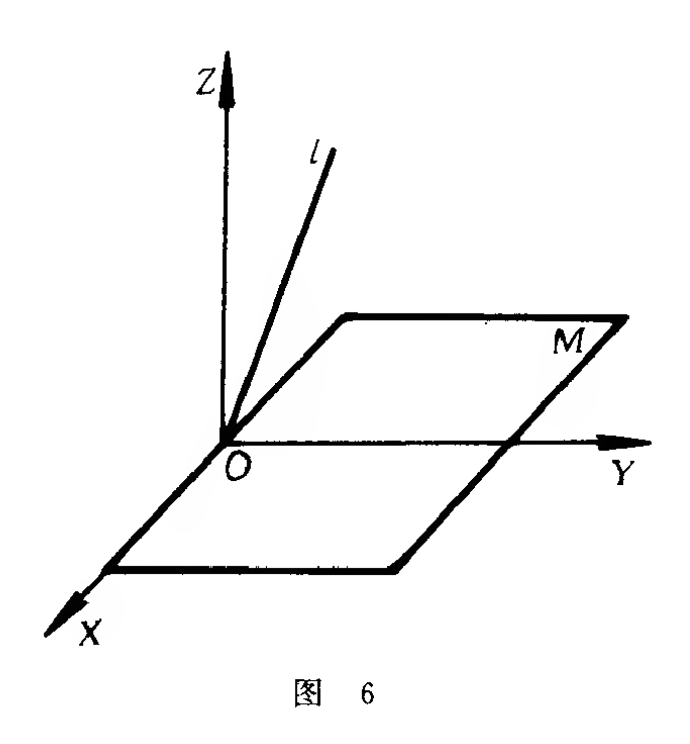


图 6

两式相减，得

$$(\alpha_1 - \alpha_1') + (\alpha_2 - \alpha_2') = 0,$$

$$\alpha_1 - \alpha_1' = \alpha_2' - \alpha_2.$$

于是 $\alpha_1-\alpha_1'\in L\cap M=\{0\}$，即 $\alpha_1-\alpha_1'=0$，$\alpha_1=\alpha_1'$. 因而又有 $\alpha_2=\alpha_2'$.

上面两个例子反映了两种情况：第一种情况向量表法不唯一；第二种情况向量表法唯一. 究其原因，主要在于第一种情况下两个子空间的交 $M_1 \cap M_2 \neq \{0\}$，而第二种情况下 $M \cap L = \{0\}$.

当我们研究一个线性空间时，常常需要设法把它分解成两个较低维的子空间 M_1, M_2（或更多个子空间 $M_1, M_2, \cdots, M_k$）的和：$V = M_1 + M_2$. 但如果出现的是第一种情况，对于我们利用这个分解式讨论问题显然是不方便的. 我们希望的是出现第二种情况. 因而，我们这一节里把第二种情况单独提出来研究.

定义 设 M_1, M_2 是线性空间 V 的两个子空间，$M_1 + M_2 = M$. 如果对 M 内任一向量 α，其表示式

$$\alpha = \alpha_1 + \alpha_2 \quad (\alpha_1 \in M_1, \alpha_2 \in M_2)$$

是唯一的，则称 M 是 M_1 与 M_2 的**直和**（亦称 $M_1 + M_2$ 是直和），记作

$$M = M_1 \oplus M_2.$$

定理 2 设 M_1, M_2 是线性空间 V 的两个有限维子空间，则下面几条互相等价：

(i) $M_1 + M_2$ 是直和；

(ii) 0 向量表法唯一，即由

$$0 = \alpha_1 + \alpha_2 \quad (\alpha_1 \in M_1, \alpha_2 \in M_2),$$

必定有 $\alpha_1 = \alpha_2 = 0$；

(iii) $M_1 \cap M_2 = \{0\}$；

(iv) $\dim M_1 + \dim M_2 = \dim(M_1 + M_2)$.

证 采用轮转方式证明这些命题等价.

(i)⇒(ii)：按定义，$M_1 + M_2$ 内任一向量表法唯一，因而零向量的表法当然是唯一的.

(ii)⇒(iii)：用反证法. 若 $M_1 \cap M_2 \neq \{0\}$，则有一 $\beta \in M_1 \cap M_2$，$\beta \neq 0$. 于是 $\beta \in M_1$，$-\beta \in M_2$，而

$$0 = \beta + (-\beta).$$

这与零向量表法唯一的假设矛盾.

(iii)⇒(iv)：利用定理 1 即得.

(iv)⇒(i)：利用定理 1 知 $\dim(M_1\cap M_2)=0$，即 $M_1\cap M_2=\{0\}$. 对任一 $\alpha\in M_1+M_2$，如果

$$\alpha=\alpha_1+\alpha_2=\alpha_1'+\alpha_2'\quad(\alpha_1,\alpha_1'\in M_1;\alpha_2,\alpha_2'\in M_2),$$

则有

$$\alpha_1-\alpha_2'=\alpha_2'-\alpha_2.$$

于是 $\alpha_1-\alpha_1'\in M_1\cap M_2=\{0\}$，即 $\alpha_1-\alpha_1'=0,\alpha_2-\alpha_2'=0$. 这说用 $\alpha_1=\alpha_1'$，$\alpha_2=\alpha_2'$，因而 α 表法唯一. ▌

为了帮助读者理解直和的概念，我们再举两个例子.

例 1 考虑线性空间 $M_n(K)$. 命 M 为 $M_n(K)$ 内主对角线上元素之和为零的方阵所成的子空间. 又命 N 为全体 n 阶数量矩阵 $dE(d\in K)$ 所成的子空间. 如果 $A\in M\cap N$，则 $A=dE$. 但 A 的主对角线上元素之和 $nd=0$，故 $d=0$. 于是 $A=0$. 由此即得 $M\cap N=\{0\}$，所以，$M+N$ 是直和.

例 2 如果在例 1 中 M 不变，而把 N 改为 $M_n(K)$ 内全体对角矩阵所组成的子空间，那么 $M+N$ 不是直和，因为

$$\begin{pmatrix}1&&&&\\&-1&&&\\&&0&&\\&&&\ddots&\\&&&&0\end{pmatrix}\in M\cap N,$$

即 $M\cap N\neq\{0\}$.

下面我们来研究多个子空间的直和. 为书写简单，我们采用如下记号

$$M_1+M_2+\cdots+M_k=\sum_{i=1}^{k}M_i.$$

定义 设 $M_1,M_2,\cdots,M_k$ 为线性空间 V 的子空间，$M_1+M_2+\cdots+M_k=M$. 如果对 M 中任一向量 α，表达式

$$\alpha=\alpha_1+\alpha_2+\cdots+\alpha_k\quad(\alpha_i\in M_i,i=1,2,\cdots,k)$$

是唯一的，则称 M 是 $M_1,M_2,\cdots,M_k$ 的**直和**（亦称 $M_1+M_2+\cdots+$

M_k 是直和)，记作

$$M = M_1 \oplus M_2 \oplus \cdots \oplus M_k.$$

定理 3 设 $M_1, M_2, \cdots, M_k$ 是线性空间 V 的有限维子空间，则下列命题互相等价：

(i) $M_1 + M_2 + \cdots + M_k$ 是直和；

(ii) 零向量的表法唯一，即若

$$0 = \alpha_1 + \alpha_2 + \cdots + \alpha_k \quad (\alpha_i \in M_i, i = 1, 2, \cdots, k),$$

则 $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$；

(iii) $M_i \cap \left(\sum\limits_{\substack{j=1 \\ j \neq i}}^{k} M_j \right) = \{0\} \quad (i = 1, 2, \cdots, k)$；

(iv) $\dim \sum\limits_{i=1}^{k} M_i = \sum\limits_{i=1}^{k} \dim M_i$.

证 采用轮转方式来证明上述命题互相等价.

(i)⇒(ii)：显然.

(ii)⇒(iii)：用反证法. 若有某个 i，使

$$M_i \cap \left(\sum_{\substack{j=1 \\ j \neq i}}^{k} M_j \right) \neq \{0\},$$

设 β 为此交集中一个非零向量，则 $-\beta \in M_i, \beta \in \sum\limits_{\substack{j=1 \\ j \neq i}}^{k} M_j$. 于是 β 可表作

$$\beta = \alpha_1 + \cdots + \alpha_{i-1} + \alpha_{i+1} + \cdots + \alpha_k \quad (\alpha_j \in M_j).$$

那么，就有

$$\begin{aligned} 0 &= \beta + (-\beta) \\ &= \alpha_1 + \cdots + \alpha_{i-1} + (-\beta) + \alpha_{i+1} + \cdots + \alpha_k, \end{aligned}$$

其中至少 $(-\beta) \neq 0$，这与零向量的表法唯一矛盾.

(iii)⇒(iv)：采用数学归纳法. $k=2$ 时即为定理 2. 今设命题对 $k-1$ 个子空间的情况成立，证明它对 k 个子空间的情况也成立.

因为 $M_1 \cap \left(\sum_{i=2}^{k} M_j\right) = \{0\}$，应用维数公式，有

$$\dim \sum_{i=1}^{k} M_i = \dim M_1 + \dim \sum_{j=2}^{k} M_j.$$

对任一 $M_j(j \geqslant 2)$ 有

$$M_j \cap \left(\sum_{\substack{s=2 \\ s \neq j}}^{k} M_s\right) \subseteq M_j \cap \left(\sum_{\substack{s=1 \\ s \neq j}}^{k} M_s\right) = \{0\}.$$

于是，按归纳假设，有

$$\dim \sum_{j=2}^{k} M_j = \sum_{j=2}^{k} \dim M_j.$$

从而

$$\dim \sum_{i=1}^{k} M_i = \sum_{i=1}^{k} \dim M_i.$$

(iv)⇒(i)：先证

$$M_i \cap \left(\sum_{\substack{j=1 \\ j \neq i}}^{k} M_j\right) = \{0\} \quad (i = 1, 2, \cdots, k).$$

对任一 i，按维数公式，有

$$\begin{aligned}
\dim \sum_{s=1}^{k} M_s &= \dim M_i + \dim \sum_{\substack{j=1 \\ j \neq i}}^{k} M_j - \dim M_i \cap \left(\sum_{\substack{j=1 \\ j \neq i}}^{k} M_j\right) \\
&\leqslant \dim M_i + \sum_{\substack{j=1 \\ j \neq i}}^{k} \dim M_j - \dim M_i \cap \left(\sum_{\substack{j=1 \\ j \neq i}}^{k} M_j\right) \\
&= \sum_{j=1}^{k} \dim M_j - \dim M_i \cap \left(\sum_{\substack{j=1 \\ j \neq i}}^{k} M_j\right).
\end{aligned}$$

由此知

$$\dim M_i \cap \left(\sum_{\substack{j=1 \\ j \neq i}}^{k} M_j\right) = 0.$$

即

$$M_i \cap \left(\sum_{\substack{j=1 \\ j \neq i}}^{k} M_j\right) = \{0\}.$$

现设 $\alpha \in \sum_{i=1}^{k} M_i$，证明 α 的表法唯一. 命

$$\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_k = \alpha_1' + \alpha_2' + \cdots + \alpha_k',$$

其中 $\alpha_i, \alpha_i' \in M_i (i = 1, 2, \cdots, k)$. 对任一 i，有

$$\alpha_i - \alpha_i' = (\alpha_1' - \alpha_1) + \cdots + (\alpha_{i-1}' - \alpha_{i-1}) + (\alpha_{i+1}' - \alpha_{i+1}) + \cdots + (\alpha_k' - \alpha_k) \in \sum_{\substack{j=1 \\ j \neq i}}^{k} M_j.$$

于是
$$\alpha_i - \alpha_i' \in M_i \cap \left(\sum_{\substack{j=1 \\ j \neq i}}^{k} M_j \right) = \{0\},$$

即 $\alpha_i - \alpha_i' = 0$，亦即 $\alpha_i = \alpha_i' (i = 1, 2, \cdots, k)$. 这说明 α 的表法是唯一的. ▌

定理 2 和定理 3 给出了两个子空间和多个子空间的和是不是直和的几种判别法则. 这些法则在许多理论问题上都是有用的.

下面再介绍一个重要的事实.

命题 4.1 设 M 是有限维线性空间 V 的一个子空间，则必存在 V 的子空间 N，使

$$V = M \oplus N.$$

证 若 $M = \{0\}$，则取 $N = V$. 下面设 $M \neq \{0\}$. 在 M 内取一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r$，则 $M = L(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r)$. 另一方面，按命题 3.2，$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r$ 可扩充成 V 的一组基

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r, \varepsilon_{r+1}, \cdots, \varepsilon_n.$$

令 $N = L(\varepsilon_{r+1}, \cdots, \varepsilon_n)$，则显见有 $V = M + N$. 又因为 $\dim V = n = r + (n - r) = \dim M + \dim N$，故由定理 2 知：$V = M \oplus N$. ▌

命题 4.1 中所指出的子空间 N 称为子空间 M 的一个**补空间**. 一个子空间的补空间不是唯一的，因为 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r$ 可以有多种方式（实际上有无穷多种方式）扩充成空间的一组基. 为了帮助读者领会这一现象，我们看一个例子.

例 3 考虑平面上以坐标原点 O 为起点的全体向量所组成的

实数域上二维线性空间 V. 过 O 点的一条直线上的全体向量组成一个一维子空间 M. 过 O 点再任意作一条不与 M 重合的直线，它上面的全体向量组成一个一维子空间 N. 现在 $M\cap N=\{0\}$，所以 $V=M\oplus N$，于是 N 是 M 的一个补空间. 显然，N 可以有无穷多种取法. 当 M,N 确定之后，从图 7 可以看出，平面上任一向量 α 可按平行四边形法则唯一地分解成

$$\alpha=\alpha_1+\alpha_2\quad(\alpha_1\in M,\alpha_2\in N).$$

如果改变 N 的选择，那么这个分解式中的 α_1,α_2 都要跟着改变.

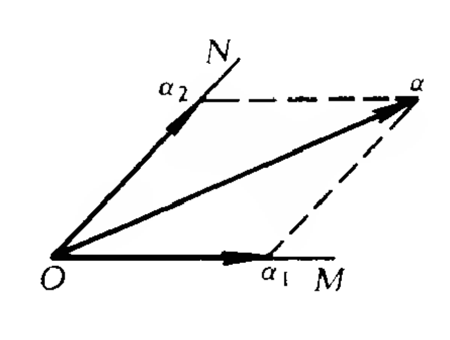


图 7

商 空 间

在这一段中，我们总假定 V 是数域 K 上的一个 n 维线性空间，M 是它的一个子空间.

现在我们利用 M 对 V 中的向量进行分类.

定义 设 α 是 V 的一个向量. 如果 V 的一个向量 α' 满足：$\alpha'-\alpha\in M$，则称 α' 与 α **模 M 同余**，记作 $\alpha'\equiv\alpha(\mathrm{mod}M)$.

不难看出，向量模 M 同余的关系具有如下三条性质：

(i) 反身性. 即 $\alpha\equiv\alpha(\mathrm{mod}M)$；

(ii) 对称性. 即：若 $\alpha'\equiv\alpha(\mathrm{mod}M)$，则 $\alpha\equiv\alpha'(\mathrm{mod}M)$；

(iii) 传递性. 即：若 $\alpha''\equiv\alpha'(\mathrm{mod}M)$，$\alpha'\equiv\alpha(\mathrm{mod}M)$，则 $\alpha''\equiv\alpha(\mathrm{mod}M)$.

这就是说，模 M 同余是 V 中的一个等价关系. 我们来把相应的等价类写出来.

设 α 是 V 中任意一个向量，定义 V 的子集

$$\alpha+M=\{\alpha+m\mid m\in M\}.$$

显然，$\alpha+M$ 内任一向量模 M 与 α 同余. 反之，模 M 与 α 同余的向量必属于 $\alpha+M$. 我们称 $\alpha+M$ 为一个模 M 的**同余类**，而 α 称为这个同余类的一个代表.

关于模 M 的同余类，有如下两条简单的事实：

(i) 若 $\alpha'\in\alpha+M$，则 $\alpha'+M=\alpha+M$. 因而，可取 $\alpha+M$ 中任一元素来作为它的代表.

这一事实是因为：$\alpha'=\alpha+m$，故 $\alpha'+M\subseteq\alpha+M$；另一方面，由 $\alpha=\alpha'-m$，又有 $\alpha+M\subseteq\alpha'+M$，从而 $\alpha'+M=\alpha+M$.

(ii) 两个模 M 同余类如不相等时就没有公共元素，即其交集为空集合.

这一事实是因为，若 $(\alpha+M)\cap(\beta+M)\neq\varnothing$，设 $\gamma\in(\alpha+M)\cap(\beta+M)$，则由(i)知 $\alpha+M=\gamma+M=\beta+M$.

线性空间 V 内每个向量必属于某一模 M 同余类(即以它自己为代表的那个同余类)，而不同的同余类不相交，于是 V 就可以看成一些彼此互相分离的同余类的并集.

我们举一个例子. 设 V 为平面上以坐标原点 O 为起点的全体向量所组成的线性空间. 命 M 为 OX 轴上全体向量组成的一维子空间. 两个平面向量 α,α' 之差 $\alpha'-\alpha\in M$ 的充分必要条件是：它们的终点落在同一条平行于 OX 轴的直线 l 上(参看图 8). 因此，现在 V 内每个模 M 的同余类可由一条平行于 OX 轴的直线来表示，该同余类是由终点落在此直线上的全体向量组成的. 显然，不同的同余类的交集是空集，而 V 则是所有这些同余类的并集(如

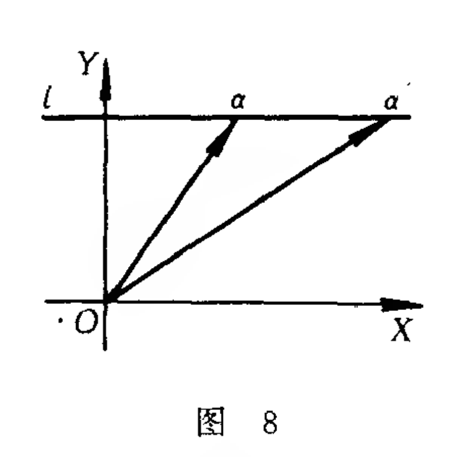


图 8

果从几何直观上看，那就是平面由平行于 OX 轴的直线并成).

命 $\overline{V}$ 表示 V 内向量模 M 的同余类的全体所成的集合(就上面的例子说，$\overline{V}$ 可以直观地理解为一条条平行于 OX 轴的直线所组成的集合). 注意 $\overline{V}$ 不是 V 的子集，它的每个元素是 V 内模 M 的一个同余类.

现在我们在集合 $\overline{V}$ 内引进加法和数乘运算：

(i) 定义

$$(\alpha + M) + (\beta + M) = (\alpha + \beta) + M;$$

(ii) 对任意 $k \in K$，定义

$$k(\alpha + M) = k\alpha + M.$$

我们证明上面的定义在逻辑上是没有矛盾的，也就是说，在同一个同余类中选择不同元素作为代表时，上面所定义的加法和数乘都不会因之而出现不同的结果. 我们分别对加法和数乘运算来做证明.

(i) 设

$$\alpha + M = \alpha' + M,\ \beta + M = \beta' + M,$$

则 $\alpha' \equiv \alpha (\mathrm{mod} M)$，$\beta' \equiv \beta (\mathrm{mod} M)$. 所以有

$$\alpha' = \alpha + m_1,\ \beta' = \beta + m_2 \quad (m_1, m_2 \in M),$$

故 $\alpha' + \beta' = \alpha + \beta + (m_1 + m_2) \in (\alpha + \beta) + M$，于是

$$(\alpha' + \beta') + M = (\alpha + \beta) + M.$$

这说明

$$(\alpha + M) + (\beta + M) = (\alpha' + M) + (\beta' + M).$$

所以，上面所定义的同余类间的加法运算不会因为各同余类代表的选择的不同而出现矛盾.

(ii) 设 $\alpha + M = \alpha' + M$，则 $\alpha' \equiv \alpha (\mathrm{mod} M)$. 所以

$$\alpha' = \alpha + m \quad (m \in M),$$

$$k\alpha' = k\alpha + km \in k\alpha + M,$$

故 $k\alpha' + M = k\alpha + M$. 于是

$$k(\alpha + M) = k(\alpha' + M).$$

所以，上面所定义的同余类的数乘运算也不会因为其代表的选择不同而出现矛盾的结果.

现在我们来证明 $\bar{V}$ 关于上面所定义的加法与数乘运算成为数域 K 上的一个线性空间. 这需要逐一检验线性空间定义中的八个条件是否满足.

(1) $$\begin{aligned}(\alpha+M)+[(\beta+M)+(\gamma+M)] &= (\alpha+M)+[(\beta+\gamma)+M] \\ &= (\alpha+\beta+\gamma)+M=[(\alpha+\beta)+M]+(\gamma+M) \\ &= [(\alpha+M)+(\beta+M)]+(\gamma+M);\end{aligned}$$

(2) $$\begin{aligned}(\alpha+M)+(\beta+M)=(\alpha+\beta)+M &= (\beta+\alpha)+M \\ &= (\beta+M)+(\alpha+M);\end{aligned}$$

(3) 以零向量为代表的同余类 $0+M=M$ 满足：

$$(\alpha+M)+(0+M)=(\alpha+0)+M=\alpha+M,$$

故 $0+M=M$ 是 $\bar{V}$ 中的零元素；

(4) $(\alpha+M)+[(-\alpha)+M]=[\alpha+(-\alpha)]+M=0+M$，即 $(-\alpha)+M$ 是 $\alpha+M$ 的负元素；

(5) $1\cdot(\alpha+M)=1\cdot\alpha+M=\alpha+M$；

(6) $$\begin{aligned}(kl)(\alpha+M)=(kl)\alpha+M &= k(l\alpha)+M \\ &= k(l\alpha+M)=k[l(\alpha+M)];\end{aligned}$$

(7) $$\begin{aligned}(k+l)(\alpha+M)=(k+l)\alpha+M &= (k\alpha+l\alpha)+M \\ &= (k\alpha+M)+(l\alpha+M) \\ &= k(\alpha+M)+l(\alpha+M);\end{aligned}$$

(8) $$\begin{aligned}k[(\alpha+M)+(\beta+M)] &= k[(\alpha+\beta)+M] \\ &= k(\alpha+\beta)+M=(k\alpha+k\beta)+M \\ &= (k\alpha+M)+(k\beta+M) \\ &= k(\alpha+M)+k(\beta+M).\end{aligned}$$

这样，$\bar{V}$ 关于所定义的加法与数乘确实组成数域 K 上的一个线性空间. 这个线性空间称为 V 对子空间 M 的**商空间**，记作 V/M.

现在考虑商空间的维数.

在 M 内取一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r$,扩充成 V 的一组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r,\varepsilon_{r+1},\varepsilon_{r+2},\cdots,\varepsilon_n.$$

我们证明:$\overline{V}$ 内的 $n-r$ 个元素

$$\varepsilon_{r+1}+M,\ \varepsilon_{r+2}+M,\ \cdots,\ \varepsilon_n+M \qquad (\text{I})$$

组成 $\overline{V}=V/M$ 的一组基.

(i) 先证(Ⅰ)在 V/M 内是线性无关的. 设

$$k_{r+1}(\varepsilon_{r+1}+M)+k_{r+2}(\varepsilon_{r+2}+M)+\cdots+k_n(\varepsilon_n+M)$$
$$=0+M.$$

于是有

$$(k_{r+1}\varepsilon_{r+1}+k_{r+2}\varepsilon_{r+2}+\cdots+k_n\varepsilon_n)+M=0+M=M,$$

即 $k_{r+1}\varepsilon_{r+1}+k_{r+2}\varepsilon_{r+2}+\cdots+k_n\varepsilon_n\in M$. 于是

$$k_{r+1}\varepsilon_{r+1}+k_{r+2}\varepsilon_{r+2}+\cdots+k_n\varepsilon_n=k_1\varepsilon_1+\cdots+k_r\varepsilon_r,$$
$$k_1\varepsilon_1+\cdots+k_r\varepsilon_r-k_{r+1}\varepsilon_{r+1}-k_{r+2}\beta_{r+2}-\cdots-k_n\varepsilon_n=0.$$

由于 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 线性无关,故 $k_{r+1}=k_{r+2}=\cdots=k_n=0$. 这证明向量组(Ⅰ)是线性无关的.

(ii) 证明任一 $\alpha+M$ 可被(Ⅰ)线性表示. 设

$$\alpha=k_1\varepsilon_1+\cdots+k_r\varepsilon_r+k_{r+1}\varepsilon_{r+1}+\cdots+k_n\varepsilon_n.$$

于是有

$$\alpha-(k_{r+1}\varepsilon_{r+1}+\cdots+k_n\varepsilon_n)=k_1\varepsilon_1+\cdots+k_r\varepsilon_r\in M.$$
$$\alpha\equiv k_{r+1}\varepsilon_{r+1}+\cdots+k_n\varepsilon_n \pmod{M}.$$

因此,有

$$\alpha+M=(k_{r+1}\varepsilon_{r+1}+\cdots+k_n\varepsilon_n)+M$$
$$=k_{r+1}(\varepsilon_{r+1}+M)+\cdots+k_n(\varepsilon_n+M).$$

综合(i)与(ii)即知向量组(I)是 V/M 的一组基,从而

$$\dim(V/M)=n-r=\dim V-\dim M.$$

习 题 四

1. 命 M 为线性空间 $M_n(K)$中全体对称矩阵所成的子空间,

而 N 为全体反对称矩阵所成的子空间，证明：

$$M_n(K)=M\oplus N.$$

2. 在线性空间 $M_n(K)$ 中，命 M,N 分别表示全体上三角、下三角矩阵所成的子空间，问是否有 $M_n(K)=M\oplus N$？为什么？

3. 设 M_1 是齐次线性方程

$$x_1+x_2+\cdots+x_n=0$$

的解空间，而 M_2 是齐次线性方程组

$$x_1=x_2=\cdots=x_n$$

的解空间，证明：$K^n=M_1\oplus M_2$.

4. 设 $V=M\oplus N$，$M=M_1\oplus M_2$，证明：

$$V=M_1\oplus M_2\oplus N.$$

5. 证明：每个 n 维线性空间都可以表示成 n 个一维子空间的直和.

6. 设 V 为数域 K 上全体 n 阶准对角矩阵

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{pmatrix}$$

（其中 A_i 为 n_i 阶方阵，$i=1,2,\cdots,s$，而 n_i 为固定的正整数）所成的线性空间，令 M_i 为全体形如

$$\begin{pmatrix} 0 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & & & \\ & & & A_i & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}$$

的准对角矩阵所成的子空间，其中 A_i 与上面矩阵中 A_i 位置相同.证明：

$$V = M_1 \oplus M_2 \oplus \cdots \oplus M_s.$$

*7. 证明子空间 M_i 的和 $\sum_{i=1}^{s} M_i$ 为直和的充分必要条件是

$$M_i \cap \left(\sum_{j=1}^{i-1} M_j \right) = \{0\} \quad (i = 2,3,\cdots,s).$$

8. 设 M,N 是线性空间 V 的两个子空间,$M \subseteq N$. 设 M 的一个补空间为 $\overline{M}$:$V = M \oplus \overline{M}$. 证明:$N = M \oplus (\overline{M} \cap N)$.

9. 设

$$V = M_1 \oplus M_2 \oplus \cdots \oplus M_s,$$

证明:在每个 $M_i(i=1,2,\cdots,s)$中各取一组基,把它们并在一起后即得 V 的一组基.

10. 设 M 为线性空间 V 的一个子空间. 在 M 内取定一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r$,用两种方式扩充为 V 的基

$$\varepsilon_1,\cdots,\varepsilon_r,\varepsilon_{r+1},\cdots,\varepsilon_n;$$

$$\varepsilon_1,\cdots,\varepsilon_r,\eta_{r+1},\cdots,\eta_n.$$

这两组基之间的过渡矩阵为 T

$$(\varepsilon_1,\cdots,\varepsilon_r,\eta_{r+1},\cdots,\eta_n) = (\varepsilon_1,\cdots,\varepsilon_r,\varepsilon_{r+1},\cdots,\varepsilon_n)T,$$

其中

$$T = \begin{pmatrix} E_r & * \\ 0 & T_0 \end{pmatrix}.$$

证明:V/M 内两组基

$$\bar{\varepsilon}_{r+1} = \varepsilon_{r+1} + M,\ \bar{\varepsilon}_{r+2} = \varepsilon_{r+2} + M,\ \cdots,\ \bar{\varepsilon}_n = \varepsilon_n + M;$$

$$\bar{\eta}_{r+1} = \eta_{r+1} + M,\ \bar{\eta}_{r+2} = \eta_{r+2} + M,\ \cdots,\ \bar{\eta}_n = \eta_n + M$$

之间的过渡矩阵为

$$(\bar{\eta}_{r+1},\cdots,\bar{\eta}_n) = (\bar{\varepsilon}_{r+1},\cdots,\bar{\varepsilon}_n)T_0.$$

第五章　线性变换

在第四章里，我们对 n 维向量空间从理论上进行了抽象化. 现在我们再来对矩阵的概念从理论上进行抽象化. 但是我们这里不讨论一般的 $m\times n$ 矩阵，而仅限于对 n 阶方阵做这个工作.

还是从线性方程组的讨论入手. 考察 n 个未知量 n 个方程的线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1,\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n.\end{cases}$$

它的系数矩阵 A 是一个 n 阶方阵. 这个方程组可以写成矩阵方程

$$AX=B.$$

在第二章中我们曾指出，上述矩阵方程与一元一次方程 $ax=b$ 很相似. 我们现在再对这两者的共同点作一些分析. 我们从另一个观点来看一元一次方程 $ax=b$：考察函数 $f(x)=ax$. x 看成是在区间 $(-\infty,+\infty)$ 内变化的一个自变量. 这个函数的定义域和值域都是 $(-\infty,+\infty)$. 在第四章 §1 开头处我们讲过，任一函数可以看成一个集合 M（即其定义域）到另一个集合 N（即其值域）的一个映射. 现在集合 M 和 N 都是区间 $(-\infty,+\infty)$. 一元一次方程的实质是：在集合 N 中给定一个元素 b，问能否在集合 M 内找到一个元素（x 的值），使它在 f 下映为 b，即 $f(x)=ax=b$. 这里所涉及的函数 $f(x)$ 是一个线性函数，其特征是

$$f(x+y)=f(x)+f(y);\quad f(kx)=kf(x).$$

对矩阵方程 $AX=B$ 可以作类似的考虑. 把 X 看作是在 K^n 内变化的 n 维向量（现在把 K^n 中的向量写成列的形式），那么，

AX 就确定出从 K^n 到 K^n 自身的一个映射，它将 X 映为 AX. 一个集合到自身的一个映射称为该集合的一个**变换**. 所以，AX 确定了 K^n 内的一个变换. 矩阵方程 $AX=B$ 的求解问题现在变成：给定 K^n 内一个向量 B，问能否在 K^n 内找到 B 的原象 X(可能不止一个)，使它们在这个变换下变到 B. 上述变换也显然具有与线性函数 $f(x)=ax$ 相类似的性质

$$A(X+Y)=AX+AY,\quad A(kX)=kAX.$$

具有这种性质的变换就称为线性变换. 因此，每个 n 阶方阵代表着 n 维向量空间 K^n 内的一个线性变换.

我们在第四章已经把 n 维向量空间抽象为一般的线性空间. 与此相应，我们应当把 n 阶方阵抽象成为线性空间中具有上述线性性质的变换. 这就是本章的出发点.

§1 线性变换的定义及运算

线性变换的定义

定义 设 V 是数域 K 上的一个线性空间，$\boldsymbol{A}$ 为 V 内的一个变换(即 V 到自身的一个映射)，满足如下条件：

(i) 对任意 $\alpha,\beta\in V$，有 $\boldsymbol{A}(\alpha+\beta)=\boldsymbol{A}\alpha+\boldsymbol{A}\beta$;

(ii) 对任意 $\alpha\in V$，任意 $k\in K$，有 $\boldsymbol{A}(k\alpha)=k\boldsymbol{A}\alpha$.

则称 $\boldsymbol{A}$ 是线性空间 V 内的一个**线性变换**.

性质(i)与(ii)与下面性质等价：

(iii) 对任意 $\alpha,\beta\in V$，任意 $k,l\in K$，有

$$\boldsymbol{A}(k\alpha+l\beta)=k\boldsymbol{A}\alpha+l\boldsymbol{A}\beta.$$

由性质(i)与(ii)推(iii)是显而易见的. 反之，设(iii)成立，则分别令 $k=l=1$ 和 $l=0$ 就得到性质(i)与(ii).

线性变换具有如下简单性质：

(1) $\boldsymbol{A}$ 把零向量变成零向量.

这只要在性质(ii)中取 $k=0$ 就可以得到了.

(2) 任给 $\alpha_1,\alpha_2,\cdots,\alpha_s\in V$，有

$$\boldsymbol{A}(k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s)=k_1\boldsymbol{A}\alpha_1+k_2\boldsymbol{A}\alpha_2+\cdots+k_s\boldsymbol{A}\alpha_s.$$

这只要连续应用性质(iii)(或(i)与(ii))就可以得到了.

(3) 若 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性相关，则 $\boldsymbol{A}\alpha_1,\boldsymbol{A}\alpha_2,\cdots,\boldsymbol{A}\alpha_s$ 线性相关.

这只要应用性质(2)，根据线性相关的定义就可以得出来了. 但要注意，相反的命题一般是不成立的，即由 $\boldsymbol{A}\alpha_1,\boldsymbol{A}\alpha_2,\cdots,\boldsymbol{A}\alpha_s$ 线性相关推不出 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性相关.

上面列举出来的性质(2)反映了线性变换的最基本的特征，在后面的讨论中常常要用到它.

在 n 维向量空间 K^n 中，一个 n 阶方阵 A 给出的变换：$X\longmapsto AX$ 是线性变换的典型例子，这在本章开头所作的分析中已经指出来了. 但上面给出的线性变换概念有更广泛的意义. 我们再来举几个例子.

例 1 设 $D_0(a,b)$ 是区间 (a,b) 内全体任意次可微的实函数 $f(x)$ 所成的集合，它关于普通函数的加法和与实数的乘法成一实数域上的线性空间. 在 $D_0(a,b)$ 内定义一个变换：

$$\boldsymbol{D}=\frac{\mathrm{d}}{\mathrm{d}x}:\ f(x)\longmapsto f'(x),$$

即让 $D_0(a,b)$ 内每个向量 $f(x)$ 在变换 $\boldsymbol{D}$ 作用下变成该函数的导函数. 这个变换就是数学分析中的求微商运算. 根据微商的性质，有

$$\boldsymbol{D}(f(x)+g(x))=\boldsymbol{D}f(x)+\boldsymbol{D}g(x);$$
$$\boldsymbol{D}(kf(x))=k\boldsymbol{D}f(x).$$

这说明求微商运算是线性空间 $D_0(a,b)$ 内的一个线性变换.

例 2 设 $C[a,b]$ 为闭区间 $[a,b]$ 上的全体连续函数所组成的实数域上的线性空间. 在 $C[a,b]$ 内定义变换如下

$$\boldsymbol{S}:\ f(x)\longmapsto\int_a^x f(t)\mathrm{d}t=F(x),$$

即让 $C[a,b]$ 内的每个向量 $f(x)$ 对应于它的变上限积分(即 $f(x)$

的一个原函数). 根据定积分的性质,我们有

$$\boldsymbol{S}(f(x)+g(x))=\boldsymbol{S}f(x)+\boldsymbol{S}g(x);$$

$$\boldsymbol{S}(kf(x))=k\boldsymbol{S}f(x).$$

这说明求变上限积分是线性空间 $C[a,b]$ 内的一个线性变换.

下面我们介绍线性空间 V 内几种特殊的线性变换.

(1) 零变换 $\mathbf{0}$.

对任意 $\alpha\in V$, $\mathbf{0}\alpha=0$. 这显然是 V 内的一个线性变换,称为**零变换**.

(2) 单位变换 $\boldsymbol{E}$.

对任意 $\alpha\in V$, $\boldsymbol{E}\alpha=\alpha$. 这显然也是 V 内的一个线性变换,称为**单位变换**或**恒等变换**.

(3) 数乘变换 $\boldsymbol{k}$.

设 k 是数域 K 内一个固定的数. 对任意 $\alpha\in V$,定义: $\boldsymbol{k}\alpha=k\alpha$. 不难验证,这也是 V 内的一个线性变换,称为**数乘变换**.

*(4) 投影变换 $\boldsymbol{P}$.

设 M 是 V 的一个子空间. 按第四章命题 4.1,存在 V 的子空间 N,使

$$V=M\oplus N.$$

于是,对任意 $\alpha\in V$,有唯一分解式

$$\alpha=\alpha_1+\alpha_2 \quad (\alpha_1\in M;\alpha_2\in N).$$

我们定义 V 内一个变换 $\boldsymbol{P}$ 如下:

$$\boldsymbol{P}\alpha=\alpha_1.$$

我们证明 $\boldsymbol{P}$ 是一个线性变换.

(i) 设 $\alpha,\beta\in V$,又

$$\alpha=\alpha_1+\alpha_2,\ \beta=\beta_1+\beta_2 \quad (\alpha_1,\beta_1\in M;\alpha_2,\beta_2\in N).$$

则

$$\alpha+\beta=(\alpha_1+\beta_1)+(\alpha_2+\beta_2) \quad (\alpha_1+\beta_1\in M;\alpha_2+\beta_2\in N).$$

按 $\boldsymbol{P}$ 的定义,有

$$\boldsymbol{P}(\alpha+\beta)=\alpha_1+\beta_1=\boldsymbol{P}\alpha+\boldsymbol{P}\beta.$$

(ii) 对任意 $k\in K$，有

$$k\alpha=k\alpha_1+k\alpha_2 \quad (k\alpha_1\in M;k\alpha_2\in N),$$

$$\boldsymbol{P}(k\alpha)=k\alpha_1=k\boldsymbol{P}\alpha.$$

线性变换 $\boldsymbol{P}$ 称为 V 对子空间 M(关于直和分解式 $V=M\oplus N$)的**投影变换**. 注意投影变换依赖于直和分解式的具体形式.

例 3 考虑平面上以坐标原点 O 为起点的全体向量所组成的实数域上二维线性空间 V. 过 O 点的直线 M 表示其某个一维子空间. 那么,任一过 O 点而又不与 M 重合的直线 N 都可以代表 M 的补空间. 此时 $V=M\oplus N$. V 对 M 关于上述直和分解式的投影变换 $\boldsymbol{P}$,就是把平面上每个向量 α 以平行于 N 的方向投影到 M 上(参看图 1).

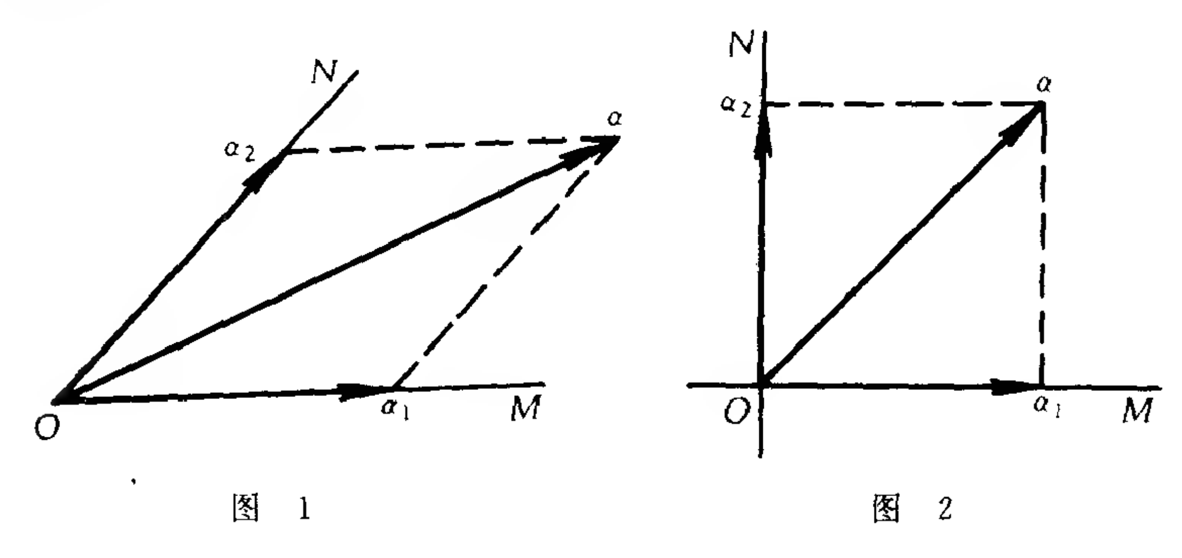


图 1 图 2

如果把 N 取在与 M 垂直的位置上,则由这样的直和分解式

$$V=M\oplus N$$

所确定的投影变换 $\boldsymbol{P}$ 称为**正投影**(参看图 2).

投影变换是一种重要的线性变换.

线性变换的运算

线性变换和矩阵一样,有加法、数乘和乘法运算.

设 V 是数域 K 上的一个线性空间,以 $L(V)$表示 V 上全体线性变换所成的集合. 现在我们在 $L(V)$内引进加法、数乘和乘法运

算如下.

1. 加法. 设 $\boldsymbol{A},\boldsymbol{B}\in L(V)$，在 V 内定义一个变换

$$(\boldsymbol{A}+\boldsymbol{B})\alpha=\boldsymbol{A}\alpha+\boldsymbol{B}\alpha.$$

那么，$\boldsymbol{A}+\boldsymbol{B}$ 也是 V 内的一个线性变换. 这是因为：

(i) $$\begin{aligned}(\boldsymbol{A}+\boldsymbol{B})(\alpha+\beta)&=\boldsymbol{A}(\alpha+\beta)+\boldsymbol{B}(\alpha+\beta)\\&=(\boldsymbol{A}\alpha+\boldsymbol{A}\beta)+(\boldsymbol{B}\alpha+\boldsymbol{B}\beta)\\&=(\boldsymbol{A}\alpha+\boldsymbol{B}\alpha)+(\boldsymbol{A}\beta+\boldsymbol{B}\beta)\\&=(\boldsymbol{A}+\boldsymbol{B})\alpha+(\boldsymbol{A}+\boldsymbol{B})\beta;\end{aligned}$$

(ii) $$\begin{aligned}(\boldsymbol{A}+\boldsymbol{B})(k\alpha)&=\boldsymbol{A}(k\alpha)+\boldsymbol{B}(k\alpha)=k\boldsymbol{A}\alpha+k\boldsymbol{B}\alpha\\&=k(\boldsymbol{A}\alpha+\boldsymbol{B}\alpha)=k(\boldsymbol{A}+\boldsymbol{B})\alpha.\end{aligned}$$

2. 数乘. 设 $\boldsymbol{A}\in L(V)$，$k\in K$，在 V 内定义一个变换

$$(k\boldsymbol{A})\alpha=k(\boldsymbol{A}\alpha).$$

请读者自行验证：$k\boldsymbol{A}$ 也是 V 内的一个线性变换.

我们指出：线性变换的加法与数乘运算满足线性空间的八条基本规律.

(1) $\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})=(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}$；

(2) $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$；

(3) $\boldsymbol{A}+\boldsymbol{0}=\boldsymbol{A}$；

(4) 对每个 $\boldsymbol{A}\in L(V)$，定义 V 内一个变换

$$(-\boldsymbol{A})\alpha=-(\boldsymbol{A}\alpha),$$

不难验证，$-\boldsymbol{A}$ 也是 V 的一个线性变换，且 $(-\boldsymbol{A})+\boldsymbol{A}=\boldsymbol{0}$；

(5) $1\cdot\boldsymbol{A}=\boldsymbol{A}$；

(6) $(kl)\boldsymbol{A}=k(l\boldsymbol{A})$；

(7) $(k+l)\boldsymbol{A}=k\boldsymbol{A}+l\boldsymbol{A}$；

(8) $k(\boldsymbol{A}+\boldsymbol{B})=k\boldsymbol{A}+k\boldsymbol{B}$.

这八条规律很容易根据定义直接加以验证，这里不详述了.

因此，$L(V)$ 关于上面定义的加法和数乘成为数域 K 上的一个线性空间.

3. 乘法. 设 $\boldsymbol{A},\boldsymbol{B}\in L(V)$，在 V 内定义一个变换

$$(\boldsymbol{AB})\alpha=\boldsymbol{A}(\boldsymbol{B}\alpha).$$

变换 $\boldsymbol{AB}$ 也是 V 内的一个线性变换，因为：

(i) $(\boldsymbol{AB})(\alpha+\beta)=\boldsymbol{A}[\boldsymbol{B}(\alpha+\beta)]=\boldsymbol{A}[\boldsymbol{B}\alpha+\boldsymbol{B}\beta]$

$$=\boldsymbol{A}(\boldsymbol{B}\alpha)+\boldsymbol{A}(\boldsymbol{B}\beta)=(\boldsymbol{AB})\alpha+(\boldsymbol{AB})\beta;$$

(ii) $(\boldsymbol{AB})(k\alpha)=\boldsymbol{A}[\boldsymbol{B}(k\alpha)]=\boldsymbol{A}[k\boldsymbol{B}\alpha]=k\boldsymbol{A}(\boldsymbol{B}\alpha)$

$$=k(\boldsymbol{AB})\alpha.$$

关于线性变换的乘法，应当指出如下几点：

(a) 乘法有结合律：$\boldsymbol{A}(\boldsymbol{BC})=(\boldsymbol{AB})\boldsymbol{C}$；

(b) 对 $k\in K$，有 $k(\boldsymbol{AB})=(k\boldsymbol{A})\boldsymbol{B}=\boldsymbol{A}(k\boldsymbol{B})$；

(c) 有左、右分配律：

$$\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{AB}+\boldsymbol{AC},$$

$$(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C}=\boldsymbol{AC}+\boldsymbol{BC};$$

(d) 单位变换 $\boldsymbol{E}$ 起着类似于 1 在数的乘法中的作用

$$\boldsymbol{AE}=\boldsymbol{EA}=\boldsymbol{A};$$

(e) $\boldsymbol{0A}=\boldsymbol{A0}=\boldsymbol{0}$；

(f) 在一般情况下，$\boldsymbol{AB}\neq\boldsymbol{BA}$，即乘法不可交换；

(g) 由 $\boldsymbol{AB}=\boldsymbol{0}$，不能推断必有 $\boldsymbol{A}=\boldsymbol{0}$ 或 $\boldsymbol{B}=\boldsymbol{0}$. 随之是由 $\boldsymbol{AB}=\boldsymbol{AC}(\boldsymbol{A}\neq\boldsymbol{0})$，不能推出 $\boldsymbol{B}=\boldsymbol{C}$.

上面几条中，读者应特别注意(f)与(g)两条. 这两条是它与数的乘法的重要的不同点.

最后，我们介绍逆变换的概念.

设 $\boldsymbol{A}\in L(V)$，如果存在 V 内的一个变换 $\boldsymbol{B}$，使

$$\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{E},$$

则称 $\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的**逆变换**. 不难看出，如果 $\boldsymbol{A}$ 的逆变换存在，那么它一定是唯一的(证法与证明逆矩阵唯一相同). 我们把 $\boldsymbol{A}$ 的逆变换记作 $\boldsymbol{A}^{-1}$. 此时 $\boldsymbol{A}$ 称为**可逆线性变换**.

在上面的定义中并没有要求 $\boldsymbol{B}$ 本身也是线性变换. 但 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 的定义仍沿用线性变换乘法的定义(事实上，V 内任意两个变换的乘法都是这样定义的，并不一定要求它们分别都是线性变

换).

我们证明:满足上述条件的 $\boldsymbol{B}$ 存在时,它一定也是 V 内的线性变换.

$$
\begin{aligned}
\text{(i)}\quad \boldsymbol{B}(\alpha+\beta)&=\boldsymbol{B}(\boldsymbol{E}\alpha+\boldsymbol{E}\beta)\\
&=\boldsymbol{B}[(\boldsymbol{AB})\alpha+(\boldsymbol{AB})\beta]\\
&=\boldsymbol{B}[\boldsymbol{A}(\boldsymbol{B}\alpha)+\boldsymbol{A}(\boldsymbol{B}\beta)]\\
&=\boldsymbol{B}[\boldsymbol{A}(\boldsymbol{B}\alpha+\boldsymbol{B}\beta)]=(\boldsymbol{BA})(\boldsymbol{B}\alpha+\boldsymbol{B}\beta)\\
&=\boldsymbol{E}(\boldsymbol{B}\alpha+\boldsymbol{B}\beta)=\boldsymbol{B}\alpha+\boldsymbol{B}\beta;
\end{aligned}
$$

$$
\begin{aligned}
\text{(ii)}\quad \boldsymbol{B}(k\alpha)&=\boldsymbol{B}(k\boldsymbol{E}\alpha)=\boldsymbol{B}[k(\boldsymbol{AB})\alpha]\\
&=\boldsymbol{B}[k\boldsymbol{A}(\boldsymbol{B}\alpha)]=\boldsymbol{B}[\boldsymbol{A}(k\boldsymbol{B}\alpha)]\\
&=(\boldsymbol{BA})(k\boldsymbol{B}\alpha)=\boldsymbol{E}(k\boldsymbol{B}\alpha)=k\boldsymbol{B}\alpha.
\end{aligned}
$$

今后,我们对线性变换也使用通常方幂的记号,即规定

$$\boldsymbol{A}^0=\boldsymbol{E},\quad \boldsymbol{A}^k=\overbrace{\boldsymbol{AA}\cdots\boldsymbol{A}}^{k\text{ 个}};$$

如果给定系数在数域 K 内的一个多项式

$$f(\lambda)=a_0\lambda^m+a_1\lambda^{m-1}+\cdots+a_{m-1}\lambda+a_m,$$

则定义

$$f(\boldsymbol{A})=a_0\boldsymbol{A}^m+a_1\boldsymbol{A}^{m-1}+\cdots+a_{m-1}\boldsymbol{A}+a_m\boldsymbol{E}.$$

另外,如果 $\boldsymbol{A}$ 可逆,则定义

$$\boldsymbol{A}^{-k}=(\boldsymbol{A}^{-1})^k.$$

习 题 一

1. 判断下面所定义的变换哪些是线性的,哪些则不是:

(1) 在线性空间 V 中,$\boldsymbol{A}\xi=\xi+\alpha$,其中 $\alpha\in V$ 是一个固定的向量;

(2) 在线性空间 V 中,令 $\boldsymbol{A}\xi=\alpha$,其中 $\alpha\in V$ 是一个固定的向量;

(3) 在 K^3 中,令 $\boldsymbol{A}(x_1,x_2,x_3)=(x_1^2,x_2+x_3,x_3^2)$;

(4) 在 K^3 中,令 $\boldsymbol{A}(x_1,x_2,x_3)=(2x_1-x_2,x_2+x_3,x_1)$;

(5) 在 $K[X]$中，令 $\boldsymbol{A}f(X)=f(X+1)$；

(6) 在 $K[X]$中，令 $\boldsymbol{A}f(X)=f(X_0)$，其中 $X_0\in K$ 是一个固定的数；

(7) 把复数域看作复数域上的线性空间，令 $\boldsymbol{A}\xi=\bar{\xi}$；

(8) 在 $M_n(K)$中，令 $\boldsymbol{A}(X)=BXC$，其中 B,C 是 K 上两个固定的 n 阶方阵.

2. 在实数域上线性空间 $D_0(a,b)$(参看本节例 1)中定义变换如下

$$\boldsymbol{A}f(x)=\frac{\mathrm{d}^2f(x)}{\mathrm{d}x^2}+x\cdot\frac{\mathrm{d}f(x)}{\mathrm{d}x}+\sin x\cdot f(x),$$

证明 $\boldsymbol{A}$ 是一个线性变换. 定义

$$\boldsymbol{B}f(x)=\left(\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right)^2+x\cdot\frac{\mathrm{d}f(x)}{\mathrm{d}x}+\sin x\cdot f(x),$$

举例说明 $\boldsymbol{B}$ 不是线性变换.

3. 在实数域上线性空间 $C[a,b]$中定义变换如下：

$$\boldsymbol{A}f(x)=\int_a^x K(t)f(t)\mathrm{d}t,$$

其中 $K(x)$是$[a,b]$上一个固定的连续函数. 证明 $\boldsymbol{A}$ 是一个线性变换.

4. 在数域 K 上全体 n 阶对称方阵所成的线性空间 V 中定义变换

$$\boldsymbol{A}X=T'XT,$$

其中 T 为一个固定的 n 阶方阵. 证明：$\boldsymbol{A}$ 是 V 中一个线性变换.

5. 在 $K[X]$中定义

$$\boldsymbol{A}f(X)=f'(X),\quad \boldsymbol{B}f(X)=Xf(X),$$

证明 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是两个线性变换，且 $\boldsymbol{AB}-\boldsymbol{BA}=\boldsymbol{E}$.

6. 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是两个线性变换，且 $\boldsymbol{AB}-\boldsymbol{BA}=\boldsymbol{E}$. 证明：对任一正整数 k，有

$$\boldsymbol{A}^k\boldsymbol{B}-\boldsymbol{B}\boldsymbol{A}^k=k\boldsymbol{A}^{k-1}.$$

7. 设线性空间 V 分解为子空间 M,N 的直和：$V=M\oplus N$. 令

$\boldsymbol{P}$ 为关于此直和分解式的 V 对 M 的投影变换，证明：

(1) $\boldsymbol{P}^2=\boldsymbol{P}$；

(2) 若 $M\neq V$，证明 $\boldsymbol{P}$ 不可逆；

(3) 命 $\boldsymbol{P}_1$ 表示 V 关于上述直和分解式对子空间 N 的投影变换，证明：$\boldsymbol{P}\boldsymbol{P}_1=\boldsymbol{P}_1\boldsymbol{P}=\boldsymbol{0}$.

*8. 设 $\boldsymbol{A}$ 是线性空间 V 中的一个线性变换，且 $\boldsymbol{A}^2=\boldsymbol{A}$. 证明：

(1) V 中任一向量 α 可分解为

$$\alpha=\alpha_1+\alpha_2,$$

其中 $\boldsymbol{A}\alpha_1=\alpha_1$，$\boldsymbol{A}\alpha_2=0$，且这种分解是唯一的；

(2) 若 $\boldsymbol{A}\alpha=-\alpha$，则 $\alpha=0$；

*9. 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是两个线性变换，$\boldsymbol{A}^2=\boldsymbol{A}$，$\boldsymbol{B}^2=\boldsymbol{B}$. 证明：若 $(\boldsymbol{A}+\boldsymbol{B})^2=\boldsymbol{A}+\boldsymbol{B}$，则 $\boldsymbol{A}\boldsymbol{B}=\boldsymbol{0}$.

10. 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是线性空间 V 的一组基，证明：线性变换 $\boldsymbol{A}$ 可逆，当且仅当 $\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n$ 线性无关.

§2 线性变换的矩阵

在第四章§2的最后一段中我们指出：对一个 n 维线性空间 V，取定一组基之后，就可以在它和 n 维向量空间 K^n 之间建立起一个同构映射. 因而，从代数学的角度可以把它们等同看待，或者换一句话说，它们具有相同的代数性质. 现在我们指出，对于 n 维线性空间 V 内全体线性变换所成的集合 $L(V)$，它和数域 K 上的 n 阶方阵的全体所成的集合 $M_n(K)$ 之间，也可以建立起类似的关系.

在这一节里，我们总假定 V 是数域 K 上的一个 n 维线性空间，$L(V)$ 是 V 内全体线性变换所成的集合.

线性变换在一组基下的矩阵

在 V 内取定一组基

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n.$$

V 内任一向量 α 可表为

$$\alpha = a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n.$$

根据 § 1 中指出的线性变换的基本性质，对于 V 内一个线性变换 $\boldsymbol{A}$，有

$$\boldsymbol{A}\alpha = a_1\boldsymbol{A}\varepsilon_1 + a_2\boldsymbol{A}\varepsilon_2 + \cdots + a_n\boldsymbol{A}\varepsilon_n.$$

这说明，只要知道了 $\boldsymbol{A}\varepsilon_1, \boldsymbol{A}\varepsilon_2 \cdots, \boldsymbol{A}\varepsilon_n$，那么任一向量 α 在 $\boldsymbol{A}$ 作用下的象 $\boldsymbol{A}\alpha$ 也就被确定了. 所以，一个线性变换决定于它对一组基的作用. 换一句话说，对 $\boldsymbol{V}$ 内两个线性变换 $\boldsymbol{A}, \boldsymbol{B}$，如果

$$\boldsymbol{A}\varepsilon_i = \boldsymbol{B}\varepsilon_i \qquad (i = 1, 2, \cdots, n),$$

那么就有 $\boldsymbol{A} = \boldsymbol{B}$（两个线性变换相等是指它们对 V 内任一向量的作用都相同）.

命题 2.1 设 V 是数域 K 上的 n 维线性空间，$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是它的一组基，则有

(i) V 内任一线性变换 $\boldsymbol{A}$，由它在 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 处的作用

$$\boldsymbol{A}\varepsilon_1, \boldsymbol{A}\varepsilon_2, \cdots, \boldsymbol{A}\varepsilon_n$$

唯一决定，即如果 $\boldsymbol{A}\varepsilon_i = \boldsymbol{B}\varepsilon_i (i = 1, 2, \cdots, n)$，则 $\boldsymbol{A} = \boldsymbol{B}$；

(ii) 任给 $\alpha_1, \alpha_2, \cdots, \alpha_n \in V$，必存在唯一的线性变换 $\boldsymbol{A}$，使

$$\boldsymbol{A}\varepsilon_i = \alpha_i \quad (i = 1, 2, \cdots, n).$$

证 现在只需证明(ii).

在 V 内定义一个变换 $\boldsymbol{A}$ 如下：若

$$\alpha = a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n,$$

则

$$\boldsymbol{A}\alpha = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n.$$

显然，现在有

$$\boldsymbol{A}\varepsilon_i = \alpha_i \quad (i = 1, 2, \cdots, n).$$

我们只要证明 $\boldsymbol{A}$ 是一个线性变换就可以了. 设

$$\beta = b_1\varepsilon_1 + b_2\varepsilon_2 + \cdots + b_n\varepsilon_n,$$

则

$$A\beta=b_1\alpha_1+b_2\alpha_2+\cdots+b_n\alpha_n.$$

而

$$\alpha+\beta=(a_1+b_1)\varepsilon_1+(a_2+b_2)\varepsilon_2+\cdots+(a_n+b_n)\varepsilon_n,$$

我们有

$$\begin{aligned}\boldsymbol{A}(\alpha+\beta)&=(a_1+b_1)\alpha_1+(a_2+b_2)\alpha_2\cdots+(a_n+b_n)\alpha_n\\&=(a_1\alpha_1+a_2\alpha_2+\cdots+a_n\alpha_n)\\&\quad+(b_1\alpha_1+b_2\alpha_2+\cdots+b_n\alpha_n)\\&=\boldsymbol{A}\alpha+\boldsymbol{A}\beta.\end{aligned}$$

同样可证：$\boldsymbol{A}(k\alpha)=k\boldsymbol{A}\alpha$，故 $\boldsymbol{A}$ 是 V 内的一个线性变换. $\boldsymbol{A}$ 的唯一性由(i)即可知. ▌

现在将 $\boldsymbol{A}\varepsilon_i$ 表为基的线性组合. 设

$$\begin{cases}\boldsymbol{A}\varepsilon_1=a_{11}\varepsilon_1+a_{21}\varepsilon_2+\cdots+a_{n1}\varepsilon_n,\\\boldsymbol{A}\varepsilon_2=a_{12}\varepsilon_1+a_{22}\varepsilon_2+\cdots+a_{n2}\varepsilon_n,\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\\boldsymbol{A}\varepsilon_n=a_{1n}\varepsilon_1+a_{2n}\varepsilon_2+\cdots+a_{nn}\varepsilon_n.\end{cases}\tag{1}$$

命

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}.$$

这就是把 $\boldsymbol{A}\varepsilon_j$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的坐标作为矩阵 A 的第 j 列($j=1,2,\cdots,n$). 因为 $\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n$ 由矩阵 A 唯一决定，从而线性变换 $\boldsymbol{A}$ 由矩阵 A 唯一决定. 我们称 n 阶方阵 A 为**线性变换 $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵**.

如果采用形式写法，(1)式可表为

$$(\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}$$

$$=(\varepsilon_1,\varepsilon_2,\cdots\varepsilon_n)A,$$

或者写成

$$\boldsymbol{A}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n)$$
$$=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A.$$

应用线性变换的矩阵可以把抽象的线性变换具体化，也就是说，可以把任一向量在线性变换下的象的坐标用原来向量的坐标表示出来.

命题 2.2 设线性变换 $\boldsymbol{A}$ 在一组基下的矩阵为 A，又设向量 α 在这组基下的坐标为

$$X=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},$$

则 $\boldsymbol{A}\alpha$ 在这组基下的坐标为 AX.

证 设这组基为 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$. 采用形式的写法，有

$$\begin{aligned}\boldsymbol{A}\alpha&=\boldsymbol{A}(x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n)\\&=\boldsymbol{A}[(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X]\\&=[\boldsymbol{A}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)]X\\&=[(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A]X\\&=(\varepsilon_1,\varepsilon_2\cdots,\varepsilon_n)(AX).\end{aligned}$$ ▌

这个命题也说明，线性空间 V 内的线性变换：$\alpha\longmapsto\boldsymbol{A}\alpha$ 相当于线性空间 K^n 中的线性变换：$X\longmapsto AX$，这正是本章开头所指出的那个线性变换.

例 1 在 $K[X]_4$ 内取定一组基

$$1,X,X^2,X^3.$$

在 $K[X]_4$ 内定义一个变换 $\boldsymbol{A}$ 如下：若

$$f(X)=a_0+a_1X+a_2X^2+a_3X^3,$$

则

$$\boldsymbol{A}f(X)=a_3+a_2X+a_1X^2+a_0X^3.$$

容易验证，$\boldsymbol{A}$ 是一个线性变换. 而因为

$$\boldsymbol{A}1=0\cdot 1+0\cdot X+0\cdot X^2+1\cdot X^3,$$
$$\boldsymbol{A}X=0\cdot 1+0\cdot X+1\cdot X^2+0\cdot X^3,$$
$$\boldsymbol{A}X^2=0\cdot 1+1\cdot X+0\cdot X^2+0\cdot X^3,$$
$$\boldsymbol{A}X^3=1+0\cdot X+0\cdot X^2+0\cdot X^3.$$

故 $\boldsymbol{A}$ 在基 $1,X,X^2,X^3$ 下的矩阵为

$$\boldsymbol{A}=\begin{pmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{pmatrix}.$$

采用形式写法，有

$$(\boldsymbol{A}1,\boldsymbol{A}X,\boldsymbol{A}X^2,\boldsymbol{A}X^3)=(1,X,X^2,X^3)A.$$

例 2 考察 $K[X]_n$ 内求微商的变换 $\boldsymbol{D}=\dfrac{\mathrm{d}}{\mathrm{d}X}$. 因为

$$\boldsymbol{D}1=0,$$
$$\boldsymbol{D}X=1,$$
$$\boldsymbol{D}X^2=2X,$$
$$\cdots\cdots\cdots\cdots$$
$$\boldsymbol{D}X^{n-1}=(n-1)X^{n-2},$$

故 $\boldsymbol{D}$ 在基 $1,X,X^2,\cdots,X^{n-1}$ 下的矩阵为

$$\boldsymbol{D}=\begin{pmatrix}0&1&0&&\\&0&2&\ddots&\\&&0&\ddots&0\\&&&\ddots&n-1\\&&&&0\end{pmatrix}.$$

由命题 2.2，对于任一

$$f(X)=a_0+a_1X+a_2X^2+\cdots+a_{n-1}X^{n-1},$$

$\boldsymbol{D}f(X)$ 在基 $1,X,X^2,\cdots,X^{n-1}$ 下的坐标应为

$$
\begin{pmatrix}
0 & 1 & 0 & & & \\
 & 0 & 2 & \ddots & & \\
 & & 0 & \ddots & 0 & \\
 & & & \ddots & n-1 & \\
 & & & & 0 &
\end{pmatrix}
\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{n-1} \end{pmatrix}
=
\begin{pmatrix} a_1 \\ 2a_2 \\ \vdots \\ (n-1)a_{n-1} \\ 0 \end{pmatrix}
$$

即 $\boldsymbol{D}f(X)=a_1+2a_2X+\cdots+(n-1)a_{n-1}X^{n-2}$. 这与多项式微商恰相符合.

例 3 考虑 K^3 中一个线性变换 $\boldsymbol{A}$. 设

$$\varepsilon_1=(1,0,0),\quad \varepsilon_2=(0,1,0),\quad \varepsilon_3=(0,0,1);$$

且

$$\boldsymbol{A}\varepsilon_1=(-1,1,0),\quad \boldsymbol{A}\varepsilon_2=(2,1,1),\quad \boldsymbol{A}\varepsilon_3=(0,-1,-1).$$

(i) 求 $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵.

因为

$$\boldsymbol{A}\varepsilon_1=-\varepsilon_1+\varepsilon_2+0\cdot\varepsilon_3,$$
$$\boldsymbol{A}\varepsilon_2=2\varepsilon_1+\varepsilon_2+\varepsilon_3,$$
$$\boldsymbol{A}\varepsilon_3=0\cdot\varepsilon_1-\varepsilon_2-\varepsilon_3,$$

故

$$(\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\boldsymbol{A}\varepsilon_3)=(\varepsilon_1,\varepsilon_2,\varepsilon_3)\begin{pmatrix} -1 & 2 & 0 \\ 1 & 1 & -1 \\ 0 & 1 & -1 \end{pmatrix}.$$

一般讲,K^n 内一个线性变换 $\boldsymbol{A}$ 在单位向量组成的基下的矩阵就是把 $\boldsymbol{A}\varepsilon_i$ 的分量作为列所排成的 n 阶方阵.

(ii) 在 K^3 中改取如下一组基

$$\eta_1=(1,1,1),\quad \eta_2=(1,1,0),\quad \eta_3(1,0,0),$$

求 $\boldsymbol{A}$ 在 η_1,η_2,η_3 下的矩阵.

现在应当求 $\boldsymbol{A}\eta_i$ 用 η_1,η_2,η_3 线性表示的系数. 因为

$$\eta_1=\varepsilon_1+\varepsilon_2+\varepsilon_3,$$
$$\eta_2=\varepsilon_1+\varepsilon_2,$$
$$\eta_3=\varepsilon_1,$$

故

$$\begin{aligned} \boldsymbol{A}\eta_1 &= \boldsymbol{A}\varepsilon_1+\boldsymbol{A}\varepsilon_2+\boldsymbol{A}\varepsilon_3 &&=(1,1,0), \\ \boldsymbol{A}\eta_2 &= \boldsymbol{A}\varepsilon_1+\varepsilon_2 &&=(1,2,1), \\ \boldsymbol{A}\eta_3 &= \boldsymbol{A}\varepsilon_1 &&=(-1,1,0). \end{aligned}$$

现在再求它们在 η_1,η_2,η_3 下的坐标. 按照第四章§2所指出的在 K^n 中求向量坐标的办法,这等价于解矩阵方程

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix}=\begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

左边的系数矩阵是以 η_1,η_2,η_3 作为列向量排成的,右边的矩阵是以 $\boldsymbol{A}\eta_1,\boldsymbol{A}\eta_2,\boldsymbol{A}\eta_3$ 作为列向量排成的. 做矩阵消元法

$$\left(\begin{array}{ccc:ccc} 1 & 1 & 1 & 1 & 1 & -1 \\ 1 & 1 & 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{array}\right)\longrightarrow\left(\begin{array}{ccc:ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & -1 & -2 \end{array}\right),$$

即得

$$(\boldsymbol{A}\eta_1,\boldsymbol{A}\eta_2,\boldsymbol{A}\eta_3)=(\eta_1,\eta_2,\eta_3)\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & -1 & -2 \end{pmatrix}.$$

这个例子告诉我们,同一个线性变换在不同的基下的矩阵一般是不相同的.

例 4 单位变换 $\boldsymbol{E}$ 在任一组基下的矩阵都是单位矩阵 E. 数乘变换 $\boldsymbol{k}$ 在任一组基下的矩阵都是数量矩阵 kE.

根据上面的讨论,在 V 中取定一组基之后,就给出了集合 $L(V)$ 到集合 $M_n(K)$ 的一个映射 σ,它把每个 $\boldsymbol{A}\in L(V)$ 映为它在该组基下的矩阵 $A\in M_n(K)$

$$\sigma:\ \boldsymbol{A}\longmapsto A.$$

命题 2.1 的(i)说明:当 $\boldsymbol{A}\neq\boldsymbol{B}$ 时,$\sigma(\boldsymbol{A})\neq\sigma(\boldsymbol{B})$. 即在同一组基下,不同的线性变换有不同的矩阵. 这表示 σ 是一个单一映射.

命题 2.1 的(ii)说明:任给 $A\in M_n(K)$,都存在一个 $\boldsymbol{A}\in L(V)$,使

$\sigma(\boldsymbol{A})=A$. 这表示 σ 是满映射.

综合以上两方面即知，σ 是 $L(V)$ 到 $M_n(K)$ 的一个一一映射.

命题 2.3 $L(V)$ 到 $M_n(K)$ 的映射 σ 有如下性质：

(i) 任给 $\boldsymbol{A},\boldsymbol{B}\in L(V)$，有 $\sigma(\boldsymbol{A}+\boldsymbol{B})=\sigma(\boldsymbol{A})+\sigma(\boldsymbol{B})$；

(ii) 任给 $\boldsymbol{A}\in L(V)$，$k\in K$，有 $\sigma(k\boldsymbol{A})=k\sigma(\boldsymbol{A})$；

(iii) 任给 $\boldsymbol{A},\boldsymbol{B}\in L(V)$，有 $\sigma(\boldsymbol{AB})=\sigma(\boldsymbol{A})\sigma(\boldsymbol{B})$.

证 (i)与(ii)留给读者作为练习，这里仅证明(iii). 证明中使用形式运算的结合律. 设

$$(\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A,$$

$$(\boldsymbol{B}\varepsilon_1,\boldsymbol{B}\varepsilon_2,\cdots,\boldsymbol{B}\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B,$$

则有

$$\begin{aligned}(\boldsymbol{AB}\varepsilon_1,\boldsymbol{AB}\varepsilon_2,\cdots,\boldsymbol{AB}\varepsilon_n)&=\boldsymbol{A}(\boldsymbol{B}\varepsilon_1,\boldsymbol{B}\varepsilon_2,\cdots,\boldsymbol{B}\varepsilon_n)\\&=\boldsymbol{A}[(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B]=[\boldsymbol{A}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)]B\\&=[\varepsilon_1,\varepsilon_1,\cdots,\varepsilon_n)A]B=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)(AB).\end{aligned}$$

这表示 $\boldsymbol{AB}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵是 AB，即

$$\sigma(\boldsymbol{AB})=\sigma(\boldsymbol{A})\sigma(\boldsymbol{B}).$$ ▌

如果不用形式运算，而是按定义具体写出各个表达式，也不难验证命题 2.3，只是比较繁琐. 读者可以自己动手推导一下.

命题 2.3 说明：映射 σ 使 $L(V)$ 和 $M_n(K)$ 之中的三种运算（加法、数乘、乘法）都有互相对应的关系，或者说 σ 保持运算. 这样的一一映射称为 $L(V)$ 到 $M_n(K)$ 的**同构映射**. 而 $L(V)$ 和 $M_n(K)$ 称为**同构**的. 从代数学的角度可以把它们等同看待，也就是说，它们具有相同的代数性质. 如果仅限于考虑加法和数乘运算，不难看出，σ 是线性空间 $L(V)$ 和 $M_n(K)$ 之间的一个同构映射.

推论 如果 $\boldsymbol{A}$ 可逆，则 $\sigma(\boldsymbol{A})=A$ 是可逆矩阵，而且

$$\sigma(\boldsymbol{A}^{-1})=A^{-1}.$$

反之，若 A 可逆，则 $\boldsymbol{A}$ 也可逆.

证 $\boldsymbol{A}$ 可逆时，有 $\boldsymbol{A}^{-1}$ 使 $\boldsymbol{AA}^{-1}=\boldsymbol{E}$. 由上面的例 4 以及命题 2.3 的(iii)，有

$$\sigma(\boldsymbol{A}A^{-1})=\sigma(\boldsymbol{A})\sigma(A^{-1})=A\sigma(A^{-1})=E=\sigma(\boldsymbol{E}).$$

故 A 可逆，且 $A^{-1}=\sigma(A^{-1})$. 反过来的命题留给读者自己证明. ▌

线性变换在不同基下的矩阵

例 3 中已经指出，同一个线性变换在不同的两组基下的矩阵一般是不一样的. 现在我们来寻找它们之间的关系.

命题 2.4 设

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n;$$

$$\eta_1,\eta_2,\cdots,\eta_n$$

是线性空间 V 的两组基，其过渡矩阵是 $T=(t_{ij})$，即

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T.$$

又设线性变换 $\boldsymbol{A}$ 在这两组基下的矩阵分别是 A 和 B，则

$$B=T^{-1}AT.$$

证 由线性变换的矩阵的定义，有

$$\boldsymbol{A}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A,$$

$$\boldsymbol{A}(\eta_1,\eta_2,\cdots,\eta_n)=(\eta_1,\eta_2,\cdots,\eta_n)B.$$

把(2)式代入上面的第二个等式，得

$$\boldsymbol{A}[(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T]=[(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T]B.$$

利用形式运算的结合律，有

$$\begin{aligned}\boldsymbol{A}[(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T]&=[\boldsymbol{A}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)]T\\&=[(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A]T\\&=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)(AT)\\&=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)(TB).\end{aligned}$$

因为 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 线性无关，故有

$$AT=TB.$$

而 T 可逆，因而有

$$B=T^{-1}AT.\ ▌$$

例 5 在例 3 中我们已经求得 K^3 中一个线性变换 $\boldsymbol{A}$ 在两组基下的矩阵

$$(A\varepsilon_1,A\varepsilon_2,A\varepsilon_3)=(\varepsilon_1,\varepsilon_2,\varepsilon_3)\begin{pmatrix}-1&2&0\\1&1&-1\\0&1&-1\end{pmatrix},$$

$$(A\eta_1,A\eta_2,A\eta_3)=(\eta_1,\eta_2,\eta_3)\begin{pmatrix}0&1&0\\1&1&1\\0&-1&-2\end{pmatrix}.$$

而

$$(\eta_1,\eta_2,\eta_3)=(\varepsilon_1,\varepsilon_2,\varepsilon_3)\begin{pmatrix}1&1&1\\1&1&0\\1&0&0\end{pmatrix},$$

不难验证，有

$$B=\begin{pmatrix}0&1&0\\1&1&1\\0&-1&-2\end{pmatrix}=\begin{pmatrix}1&1&1\\1&1&0\\1&0&0\end{pmatrix}^{-1}\begin{pmatrix}-1&2&0\\1&1&-1\\0&1&-1\end{pmatrix}\begin{pmatrix}1&1&1\\1&1&0\\1&0&0\end{pmatrix}.$$

定义 对数域 K 上的两个 n 阶方阵 A 与 B，如果存在 K 上一个 n 阶可逆的方阵 T，使 $B=T^{-1}AT$，则称 B 与 A **相似**，记作 $B\sim A$.

矩阵的相似关系具有如下性质：

(1) 反身性：$A\sim A$. 这是因为 $A=E^{-1}AE$；

(2) 对称性：若 $B\sim A$，则 $A\sim B$. 这是因为当 $B=T^{-1}AT$ 时，有 $A=(T^{-1})^{-1}BT^{-1}$；

(3) 传递性：若 $A\sim B$，$B\sim C$，则 $A\sim C$. 这是因为当 $A=T_1^{-1}BT_1$，$B=T_2^{-1}CT_2$ 时，有

$$A=T_1^{-1}(T_2^{-1}CT_2)T_1=(T_2T_1)^{-1}C(T_2T_1).$$

这说明矩阵的相似关系是一个等价关系. 我们把数域 K 上全体 n 阶方阵的集合 $M_n(K)$ 在相似关系下的等价类称作相似类. 于是 $M_n(K)$ 可分解为互不相交的相似类的并.

由于上述性质，我们可以把 K 上 n 阶方阵的集合 $M_n(K)$ 中的元素按相似关系进行分类，凡是互相之间存在相似关系的矩阵属

于同一类，不同的相似类之间没有公共元素(交是空集). 下面一个命题阐明了相似类的实际意义.

命题 2.5 两个 n 阶方阵 A,B 相似的充分必要条件是，它们是 V 内某一线性变换 $\boldsymbol{A}$ 在不同基下的矩阵.

证 充分性已在前面阐述了，现在我们证明必要性.

按命题 2.1 的(ii)，我们可以找到 V 内一个线性变换 $\boldsymbol{A}$，使它在 V 的某一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵为 A. 因为 $B\sim A$，故存在可逆矩阵 T，使 $B=T^{-1}AT$. 命

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T,$$

由第四章命题 2.2 的(ii)可知，$\eta_1,\eta_2,\cdots,\eta_n$ 是 V 的一组基. 根据前面的推理，$\boldsymbol{A}$ 在这组新基下的矩阵为 $T^{-1}AT=B$. ▌

由此可知，$M_n(K)$内每一个相似类实际上不过是同一个线性变换 $\boldsymbol{A}$ 在不同基下的矩阵而已. 从这一认识出发，自然就会提出这样的问题：能不能设法在 V 中找出一组基，使 $\boldsymbol{A}$ 在这组基下的矩阵具有最简单的形式呢？或者换一句话说，能不能在 $M_n(K)$内的每一个相似类里去找出一个形式最为简单的矩阵来作为该相似类的代表呢?这就是矩阵在相似关系下的标准形问题. 本章后面的几节，主要就是讨论这个问题的.

习　题　二

1. 在 K^4 内找出一个线性变换 $\boldsymbol{A}$，使 $\boldsymbol{A}$ 在基

$$\varepsilon_1=(1,2,-1,0),\quad \varepsilon_2=(1,-1,1,1),$$
$$\varepsilon_3=(-1,2,1,1),\quad \varepsilon_4=(-1,-1,0,1)$$

下有矩阵

$$A=\begin{pmatrix}-1&1&0&1\\1&1&1&0\\0&0&1&0\\-1&2&1&3\end{pmatrix}.$$

2. 求下列线性变换在指定基下的矩阵：

(1) 在 K^3 中，$\boldsymbol{A}(x_1,x_2,x_3)=(2x_1-x_2,x_2+x_3,x_1)$，而基取：$\varepsilon_1=(1,0,0),\varepsilon_2=(0,1,0),\varepsilon_3=(0,0,1)$.

(2) 在平面直角坐标系内，取 $\varepsilon_1,\varepsilon_2$ 为两个坐标向量. 令 $\boldsymbol{A}$ 为平面上向量对第一和第三象限分角线的垂直投影，$\boldsymbol{B}$ 是平面上向量对 $L(\varepsilon_2)$的垂直投影，求 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{AB}$ 在 $\varepsilon_1,\varepsilon_2$ 下的矩阵.

(3) 在 $K[X]_n$ 中，令

$$\boldsymbol{A}f(X)=f(X+1)-f(X),$$

基取为

$$\varepsilon_0=1,\varepsilon_i=\frac{X(X-1)\cdots(X-i+1)}{i!}\quad(i=1,2,\cdots,n-1).$$

(4) 在 $D_0(a,b)$中取 6 个线性无关向量

$$\varepsilon_1=e^{\alpha x}\cos\beta x,\qquad \varepsilon_2=e^{\alpha x}\sin\beta x,$$
$$\varepsilon_3=xe^{\alpha x}\cos\beta x,\qquad \varepsilon_4=xe^{\alpha x}\sin\beta x,$$
$$\varepsilon_5=\frac{1}{2}x^2e^{\alpha x}\cos\beta x,\quad \varepsilon_6=\frac{1}{2}x^2e^{\alpha x}\sin\beta x.$$

令 $V=L(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_6)$. 定义

$$\boldsymbol{D}f(x)=f'(x).$$

证明 $\boldsymbol{D}$ 是 V 内的一个线性变换，并求 $\boldsymbol{D}$ 在 V 的上述一组基下的矩阵.

(5) 已知 K^3 中线性变换 $\boldsymbol{A}$ 在基

$$\eta_1=(-1,1,1),\ \eta_2=(1,0,-1),\ \eta_3=(0,1,1)$$

下的矩阵是

$$A=\begin{pmatrix}1&0&1\\1&1&0\\-1&2&1\end{pmatrix},$$

求 $\boldsymbol{A}$ 在基

$$\varepsilon_1=(1,0,0),\ \varepsilon_2=(0,1,0),\ \varepsilon_3=(0,0,1)$$

下的矩阵.

(6) 在 K^3 中定义线性变换 $\boldsymbol{A}$ 如下：

$$\boldsymbol{A}\eta_1=(-5,0,3),\qquad \eta_1=(-1,0,2),$$
$$\boldsymbol{A}\eta_2=(0,-1,6),\qquad \eta_2=(0,1,1),$$
$$\boldsymbol{A}\eta_3=(-5,-1,9).\quad \eta_3=(3,-1,0).$$

求 $\boldsymbol{A}$ 在基

$$\varepsilon_1=(1,0,0),\ \varepsilon_2=(0,1,0),\ \varepsilon_3=(0,0,1)$$

下的矩阵.

(7) 接上. 求 $\boldsymbol{A}$ 在 η_1,η_2,η_3 下的矩阵.

3. 在 $M_2(K)$中定义变换如下:

$$\boldsymbol{A}X=AX-XA,\quad X\in M_2(K),$$

其中 A 是 K 上一个固定的二阶方阵. 证明:

(1) $\boldsymbol{A}$ 是 $M_2(K)$内的一个线性变换;

(2) 在 $M_2(K)$中取一组基

$$\varepsilon_{11}=\begin{pmatrix}1&0\\0&0\end{pmatrix},\ \varepsilon_{12}=\begin{pmatrix}0&1\\0&0\end{pmatrix},\ \varepsilon_{21}=\begin{pmatrix}0&0\\1&0\end{pmatrix},\ \varepsilon_{22}=\begin{pmatrix}0&0\\0&1\end{pmatrix}.$$

求 $\boldsymbol{A}$ 在这组基下的矩阵.

4. 在 K^n 中取 n 个单位向量(以列的形式表示)

$$\varepsilon_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},\quad \varepsilon_2=\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\ \cdots,\ \varepsilon_n=\begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}.$$

对一个 n 阶方阵 A,定义 K^n 中的线性变换

$$\boldsymbol{A}X=AX\quad (X\in K^n),$$

求 $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵.

5. 设三维线性空间 V 内一个线性变换 $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵为

$$A=\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}.$$

（1）求 $\boldsymbol{A}$ 在基 $\varepsilon_3,\varepsilon_2,\varepsilon_1$ 下的矩阵；

（2）求 $\boldsymbol{A}$ 在基 $\varepsilon_1,k\varepsilon_2,\varepsilon_3$ 下的矩阵（$k\neq 0$）；

（3）求 $\boldsymbol{A}$ 在基 $\varepsilon_1+\varepsilon_2,\varepsilon_2,\varepsilon_3$ 下的矩阵.

6. 设 $\boldsymbol{A}$ 是线性空间 V 内的线性变换. 如果 $\boldsymbol{A}^{k-1}\xi\neq 0$，但 $\boldsymbol{A}^k\xi=0$，求证：$\xi,\boldsymbol{A}\xi,\cdots,\boldsymbol{A}^{k-1}\xi(k>0)$ 线性无关.

7. 在 n 维线性空间中，设有线性变换 $\boldsymbol{A}$ 与向量 ξ，使 $\boldsymbol{A}^{n-1}\xi\neq 0$，但 $\boldsymbol{A}^n\xi=0$. 求证：$\boldsymbol{A}$ 在某一组基下的矩阵是

$$\begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}.$$

8. 设四维线性空间 V 内一个线性变换 $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 下的矩阵为

$$\begin{pmatrix} 1 & 0 & 2 & 1 \\ -1 & 2 & 1 & 3 \\ 1 & 2 & 5 & 5 \\ 2 & -2 & 1 & -2 \end{pmatrix}.$$

求 $\boldsymbol{A}$ 在 $\eta_1=\varepsilon_1-2\varepsilon_2+\varepsilon_4,\eta_2=3\varepsilon_2-\varepsilon_3-\varepsilon_4,\eta_3=\varepsilon_3+\varepsilon_4,\eta_4=2\varepsilon_4$ 下的矩阵.

9. 在 K^4 内一个线性变换 $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 下的矩阵为

$$\begin{pmatrix} 1 & -1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & 1 \end{pmatrix},$$

求它在基 $\eta_1,\eta_2,\eta_3,\eta_4$ 下的矩阵. 其中

（1）$\varepsilon_1=(1,2,-1,0)$，　　$\eta_1=(2,1,0,1)$，

$\varepsilon_2=(1,-1,1,1)$，　　$\eta_2=(0,1,2,2,)$，

$\varepsilon_3=(-1,2,1,1)$，　　$\eta_3=(-2,1,1,2)$，

$\varepsilon_4=(-1,-1,0,1)$.　　$\eta_4=(1,3,1,2)$.

(2) $\varepsilon_1=(1,1,1,1)$, $\eta_1=(1,1,0,1)$,

$\varepsilon_2=(1,1,-1,-1)$, $\eta_2=(2,1,3,1)$,

$\varepsilon_3=(1,-1,1,-1)$, $\eta_3=(1,1,0,0)$,

$\varepsilon_4=(1,-1,-1,1)$. $\eta_4=(0,1,-1,-1)$.

10. 在 K^3 中给定两组基

$\varepsilon_1=(1,0,1)$, $\eta_1=(1,2,-1)$,

$\varepsilon_2=(2,1,0)$, $\eta_2=(2,2,-1)$,

$\varepsilon_3=(1,1,1)$. $\eta_3=(2,-1,-1)$.

定义线性变换

$$\boldsymbol{A}\varepsilon_i=\eta_i \quad (i=1,2,3).$$

(1) 求 $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵；

(2) 求 $\boldsymbol{A}$ 在基 η_1,η_2,η_3 下的矩阵.

11. 证明对角矩阵

$$\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix} \quad 与 \quad \begin{pmatrix}\lambda_{i_1} & & & \\ & \lambda_{i_2} & & \\ & & \ddots & \\ & & & \lambda_{i_n}\end{pmatrix}$$

相似，其中 $i_1,i_2,\cdots,i_n$ 是 $1,2,\cdots,n$ 的一个排列.

12. 若 A 可逆，证明 AB 与 BA 相似.

13. 若 A 与 B 相似，C 与 D 相似，证明：

$$\begin{pmatrix}A & 0\\ 0 & C\end{pmatrix} \quad 与 \quad \begin{pmatrix}B & 0\\ 0 & D\end{pmatrix}$$

相似.

14. 设 V 是数域 K 上的 n 维线性空间，证明：

(1) V 内全体线性变换所成的 K 上线性空间 $L(V)$ 的维数等于 n^2；

(2) 对 V 内任一线性变换 $\boldsymbol{A}$，存在一个次数 $\leqslant n^2$ 的多项式 $f(\lambda)$（系数在 K 内），使 $f(\boldsymbol{A})=\boldsymbol{0}$.

15. 设 B 与 A 相似，$f(\lambda)$ 为任一多项式，证明 $f(B)$ 与 $f(A)$

相似.

§3 特征值与特征向量

对 n 维线性空间 V 内的一个线性变换 $\boldsymbol{A}$,我们希望能找到一组基

$$\eta_1,\eta_2,\cdots,\eta_n,$$

使 $\boldsymbol{A}$ 在这一组基下的矩阵具有最简单的形式. 在第二章中我们已经知道,对于矩阵运算来说,对角形最为简单. 因此,自然要问:有没有可能找到一组基,使 $\boldsymbol{A}$ 在这组基下的矩阵具有对角形? 亦即

$$(\boldsymbol{A}\eta_1,\boldsymbol{A}\eta_2,\cdots,\boldsymbol{A}\eta_n)=(\eta_1,\eta_2,\cdots,\eta_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}.$$

把上面的关系式具体写出来,就是

$$\begin{aligned}\boldsymbol{A}\eta_1&=\lambda_1\eta_1,\\ \boldsymbol{A}\eta_2&=\qquad\lambda_2\eta_2,\\ &\cdots\cdots\cdots\cdots\cdots\cdots\\ \boldsymbol{A}\eta_n&=\qquad\qquad\lambda_n\eta_n.\end{aligned}$$

经过较深入的研究之后就可以知道,这并不是总能办到的. 但上面的分析却给了我们一个重要的启示,即研究一个线性变换 $\boldsymbol{A}$,很重要的是去寻找满足条件:$\boldsymbol{A}\xi=\lambda\xi$ 的数 λ 和非零向量 ξ. 这一点就是本节的中心内容.

特征值与特征向量的定义

定义 设 V 是数域 K 上的一个线性空间,$\boldsymbol{A}$ 是 V 内一个线性变换. 如果对 K 内一个数 λ,存在 V 的一个向量 $\xi\neq0$,使

$$\boldsymbol{A}\xi=\lambda\xi,$$

则称 λ 为 $\boldsymbol{A}$ 的一个**特征值**,而 ξ 称为属于特征值 λ 的**特征向量**.

这里要注意两点:

(i) 特征向量 ξ 一定要是非零向量；

(ii) λ 必需属于数域 K，否则数乘 $\lambda\xi$ 没有意义.

现在设 λ 是 $\boldsymbol{A}$ 的一个特征值，那么，我们有如下两条简单的事实：

(1) 如 ξ_1,ξ_2 是 $\boldsymbol{A}$ 的属于特征值 λ 的两个特征向量，则当 $\xi_1+\xi_2\neq 0$ 时，$\xi_1+\xi_2$ 也是 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量. 这是因为

$$\boldsymbol{A}(\xi_1+\xi_2)=\boldsymbol{A}\xi_1+\boldsymbol{A}\xi_2=\lambda\xi_1+\lambda\xi_2=\lambda(\xi_1+\xi_2).$$

(2) 若 ξ 是 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量，则对任一 $k\in K$，$k\neq 0$，$k\xi$ 也是 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量. 这是因为：$k\xi\neq 0$，且

$$\boldsymbol{A}(k\xi)=k\boldsymbol{A}\xi=k(\lambda\xi)=\lambda(k\xi).$$

从(1)与(2)可知：若把 $\boldsymbol{A}$ 的属于特征值 λ 的全部特征向量再加上零向量(我们有 $\boldsymbol{A}0=\lambda\cdot 0$)合并成 V 的一个子集，那么，它对加法和数乘封闭. 按第四章命题 3.1，它是 V 的一个子空间. 将这个子空间记为 V_λ，采用集合论的记号，可表示为

$$V_\lambda=\{\xi\in V\mid \boldsymbol{A}\xi=\lambda\xi\}.$$

定义 V_λ 称为线性变换 $\boldsymbol{A}$ 的属于特征值 λ 的**特征子空间**.

当 λ 是 $\boldsymbol{A}$ 的一个特征值时，V_λ 不是零子空间，其中必定包含有无穷多个向量. 所以，$\boldsymbol{A}$ 的属于特征值 λ 的特征向量有无穷多个. 但我们只要求出 V_λ 的一组基，也就等于求出它们的全体了.

特征值与特征向量的计算法

设 V 是数域 K 上的 n 维线性空间，$\boldsymbol{A}$ 是 V 内的一个线性变换. 我们研究：

(1) 如何求出 $\boldsymbol{A}$ 的全部特征值？

(2) 对 $\boldsymbol{A}$ 的每个特征值 λ，如何求出属于它的全部特征向量？这个问题等价于寻求特征子空间 V_λ 的一组基.

在 V 内取定一组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n.$$

设 $\mathcal{A}$ 在此组基下的矩阵为 $A=(a_{ij})$，即

$$(\mathcal{A}\varepsilon_1,\mathcal{A}\varepsilon_2,\cdots,\mathcal{A}\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A.$$

设 $\lambda\in K$ 是 $\mathcal{A}$ 的一个特征值，ξ 是属于特征值 λ 的一个特征向量，那么，有

$$\mathcal{A}\xi=\lambda\xi.$$

现设 ξ 在这组基下的坐标为

$$X=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\neq 0.$$

由命题 2.2，$\mathcal{A}\xi$ 在这组基下的坐标为 AX. 而 $\lambda\xi$ 的坐标是 λX，于是

$$AX=\lambda X.$$

具体写出，就是

$$\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\lambda\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}.$$

移项，得

$$\lambda X-AX=0,$$

最后得到

$$(\lambda E-A)X=0. \tag{1}$$

因此，X 满足一个齐次线性方程，其系数矩阵为

$$\lambda E-A=\begin{pmatrix}\lambda-a_{11}&-a_{12}&\cdots&-a_{1n}\\-a_{21}&\lambda-a_{22}&\cdots&-a_{2n}\\\vdots&\vdots&&\vdots\\-a_{n1}&-a_{n2}&\cdots&\lambda-a_{nn}\end{pmatrix}.$$

根据第三章定理 1（参看第三章 §2），齐次线性方程组(1)有非零解的充分必要条件是：其系数矩阵 $\lambda E-A$ 的行列式为零，即

$$|\lambda E-A|=0. \tag{2}$$

具体写出，就是

$$\begin{vmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{vmatrix}=0.$$

综合上面的分析，不难得出如下两条结论：

1. $\lambda\in K$ 是 A 的特征值的充分必要条件是它满足方程式(2)；

2. 对于特征值 λ，V_λ 由坐标(在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下)满足齐次线性方程组(1)的全体向量组成(注意这时齐次线性方程组的系数矩阵属于 $M_n(K)$，它的解可以限制在数域 K 内). 而 V_λ 的一组基可由齐次线性方程组(1)的一个基础解系来给出.

我们在上面所作的分析给出了以上两条结论的必要性的证明，而其充分性是比较明显的，请读者自行证明.

这样，本段开头所提出的两个问题已经得到解答.

行列式 $|\lambda E-A|$ 的元素是常数或 λ 的一次多项式. 根据行列式的定义(或完全展开式)不难看出它应当是 λ 的多项式.

定义 对一个 n 阶方阵 $A=(a_{ij})$，令

$$f(\lambda)=|\lambda E-A|=\begin{vmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{vmatrix}.$$

称 $f(\lambda)$ 为矩阵 A 的**特征多项式**. $f(\lambda)$ 的根称为矩阵 A 的**特征值**或**特征根**.

设 λ_0 是矩阵 A 的一个特征值，而 X 是一个非零列向量，使 $AX=\lambda_0X$. 也就是说，X 是齐次线性方程组

$$(\lambda_0E-A)X=0$$

的一个非零解向量，我们就称 X 是矩阵 A 的属于特征值 λ_0 的**特征向量**.

现在要求线性变换 **A** 的全部特征值，只要先求它在某一组基下的矩阵 A 的全部特征值，然后从中挑出属于数域 K 的那一些就是了. 这里要注意矩阵 A 的特征值和线性变换 **A** 的特征值是有所不同的. A 的特征值是 $f(\lambda)=0$ 的全部根，其中有些可能不在数域 K 内，而线性变换 **A** 的特征值仅限于属于数域 K 的那一些.

现在我们把计算线性变换 **A** 的特征值和特征向量的步骤归纳如下：

(1) 在 V 中给定一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，求 **A** 在这组基下的矩阵 A.

(2) 计算特征多项式 $f(\lambda)=|\lambda E-A|$.

(3) 求 $f(\lambda)=0$ 的属于数域 K 的那些根

$$\lambda_1,\lambda_2,\cdots,\lambda_s.$$

(4) 对每个 $\lambda_i(i=1,2,\cdots,s)$ 求齐次线性方程组

$$(\lambda_i E-A)X=0$$

的一个基础解系. 这个齐次线性方程组具体写出来就是

$$\begin{pmatrix} \lambda_i-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda_i-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda_i-a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=0.$$

注意其中的 λ_i 是在(3)中求出的，是已知数，不是未知量.

(5) 以(4)中求出的基础解系为坐标写出 V 中一个向量组，它就是 V_{λ_i} 的一组基.

例 1 设三维线性空间 V 内一个线性变换 **A** 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵为

$$A=\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix},$$

求 **A** 的全部特征值和对应的特征向量.

解 本例中 **A** 的矩阵已给出. 下面分两步计算：

(i) 求特征多项式和特征根.

$$f(\lambda)=|\lambda E-A|=\begin{vmatrix}\lambda-1 & -2 & -2\\ -2 & \lambda-1 & -2\\ -2 & -2 & \lambda-1\end{vmatrix}$$

$$=\begin{vmatrix}\lambda-5 & -2 & -2\\ \lambda-5 & \lambda-1 & -2\\ \lambda-5 & -2 & \lambda-1\end{vmatrix}=(\lambda-5)\begin{vmatrix}1 & -2 & -2\\ 1 & \lambda-1 & -2\\ 1 & -2 & \lambda-1\end{vmatrix}$$

$$=(\lambda-5)\begin{vmatrix}1 & -2 & -2\\ 0 & \lambda+1 & 0\\ 0 & 0 & \lambda+1\end{vmatrix}=(\lambda-5)(\lambda+1)^2.$$

$f(\lambda)$的根为 $\lambda_1=5,\lambda_2=-1$(二重根). 因为整数必属于任一数域 K,所以 λ_1,λ_2 均为 A 的特征值.

(ii) 求每个特征值对应的特征向量.

$\lambda_1=5$ 时,解以 $\lambda_1 E-A=5E-A$ 为系数矩阵的齐次线性方程组. 采用矩阵消元法

$$\lambda_1 E-A=\begin{pmatrix}4 & -2 & -2\\ -2 & 4 & -2\\ -2 & -2 & 4\end{pmatrix}\longrightarrow\begin{pmatrix}-2 & -2 & 4\\ -2 & 4 & -2\\ 4 & -2 & -2\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1 & 1 & -2\\ -1 & 2 & -1\\ 2 & -1 & -1\end{pmatrix}\longrightarrow\begin{pmatrix}1 & 1 & -2\\ 0 & 3 & -3\\ 0 & -3 & 3\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1 & 1 & -2\\ 0 & 1 & -1\\ 0 & 0 & 0\end{pmatrix}$$

$$\Longleftrightarrow\begin{cases}x_1+x_2-2x_3=0,\\ \quad\ \ x_2-\ x_3=0.\end{cases}$$

移项,得

$$\begin{cases}x_1+x_2=2x_3,\\ \quad\ \ x_2=x_3.\end{cases}$$

令 $x_3=1$,得基础解系 $\eta_1=(1,1,1)$,它对应于 A 的特征向量

$\varepsilon_1+\varepsilon_2+\varepsilon_3$，于是它是特征子空间 V_{λ_1} 的一组基，即

$$V_{\lambda_1}=L(\varepsilon_1+\varepsilon_2+\varepsilon_3).$$

$\lambda_2=-1$ 时，

$$\lambda_2E-A=\begin{pmatrix}-2&-2&-2\\-2&-2&-2\\-2&-2&-2\end{pmatrix}\longrightarrow\begin{pmatrix}1&1&1\\0&0&0\\0&0&0\end{pmatrix}$$

$$\Longleftrightarrow x_1+x_2+x_3=0.$$

移项，得

$$x_1=-x_2-x_3.$$

取 $x_2=1,x_3=0$ 得 $\eta_1=(-1,1,0)$；取 $x_2=0,x_3=1$，得 $\eta_2=(-1,0,1)$. 这个基础解系对应于 $\boldsymbol{A}$ 的一个特征向量组：$-\varepsilon_1+\varepsilon_2,-\varepsilon_1+\varepsilon_3$，它们构成 V_{λ_2} 的一组基，即

$$V_{\lambda_2}=L(-\varepsilon_1+\varepsilon_2,-\varepsilon_1+\varepsilon_3).$$

例 2 在线性空间 $K[X]_n$ 中取一组基

$$1,X,\frac{1}{2!}X^2,\cdots,\frac{1}{(n-1)!}X^{n-1}.$$

容易求出微商变换 $\boldsymbol{D}f(X)=f'(X)$ 在这组基下的矩阵为

$$D=\begin{pmatrix}0&1&0&&\\&0&1&\ddots&\\&&\ddots&\ddots&0\\&&&\ddots&1\\&&&&0\end{pmatrix}.$$

(i) D 的特征多项式为

$$f(\lambda)=|\lambda E-D|=\begin{vmatrix}\lambda&-1&&\\&\lambda&\ddots&\\&&\ddots&-1\\&&&\lambda\end{vmatrix}=\lambda^n.$$

它的特征根仅有一个（n 重根），即 $\lambda_1=0\in K$. 故 $\boldsymbol{D}$ 仅有一个特征值 $\lambda_1=0$.

(ii) 求 $\lambda_1=0$ 对应的特征向量.

$$\lambda_1 E-D=-D=\begin{pmatrix}0 & -1 & & \\ & 0 & \ddots & \\ & & \ddots & -1 \\ & & & 0\end{pmatrix}\longrightarrow\begin{pmatrix}0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0\end{pmatrix}$$

$$\Longleftrightarrow\begin{cases}x_2 & =0, \\ \quad x_3 & =0, \\ \quad\quad \cdots\cdots\cdots\cdots & \\ \quad\quad\quad x_n=0. & \end{cases}$$

在这个齐次线性方程组中，仅有 x_1 是自由未知量，取 $x_1=1$，得基础解系 $\eta_1=(1,0,\cdots,0)$，它对应于 $\boldsymbol{D}$ 的特征向量

$$x_1\cdot 1+x_2X+\cdots+x_nX^{n-1}=1.$$

于是 $V_{\lambda_1}=L(1)$，即 $\boldsymbol{D}$ 的属于特征值 0 的特征向量为任一非零的常数. 这与数学分析中的结论一致.

例 3 平面上全体向量组成实数域上一个二维线性空间. 取直角坐标系的坐标向量 $\varepsilon_1,\varepsilon_2$ 作为它的一组基. 设线性变换 $\boldsymbol{A}$ 在此基下的矩阵为

$$A=\begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}\quad(\theta\neq k\pi).$$

A 的特征多项式为

$$f(\lambda)=|\lambda E-A|=\begin{vmatrix}\lambda-\cos\theta & \sin\theta \\ -\sin\theta & \lambda-\cos\theta\end{vmatrix}$$
$$=\lambda^2-2\lambda\cos\theta+1.$$

因 $\theta\neq k\pi$，这个二次方程仅有复数根，即矩阵 A 的特征值都是复数. 但我们现在考虑的是实数域上的线性空间，故 $\boldsymbol{A}$ 没有特征值，因而也没有特征向量.

从解析几何的知识可知，$\boldsymbol{A}$ 代表的是平面绕坐标原点 O 旋转 θ 角的变换. 从几何直观即可看出，当 $\theta\neq k\pi$ 时，平面上不存在非零向量 ξ，满足 $\boldsymbol{A}\xi=\lambda\xi$（即 ξ 旋转 θ 角后仍落在原向量所在的直线上）.

*特征多项式的基本性质

命题 3.1 相似的矩阵有相同的特征多项式.

证 设 $B=T^{-1}AT$,则

$$
\begin{aligned}
|\lambda E-B| &= |\lambda E-T^{-1}AT| = |T^{-1}(\lambda E-A)T| \\
&= |T^{-1}||\lambda E-A||T| = |\lambda E-A||T^{-1}T| \\
&= |\lambda E-A|\cdot|E| = |\lambda E-A|. \blacksquare
\end{aligned}
$$

因为一个线性变换 $\boldsymbol{A}$ 在不同基下的矩阵是相似的,根据上述命题,它们的特征多项式相同.因而,我们把 $\boldsymbol{A}$ 在任一组基下的矩阵的特征多项式称为 $\boldsymbol{A}$ 的**特征多项式**.这个命题又从理论上指明:在用前面讲的办法计算线性变换的特征值和特征向量时,不会因为所选择的基不相同而得到不同的结果.

现在我们利用第三章§3 中的行列式的完全展开公式来对矩阵 A 的特征多项式作一些初步的探讨.我们有:

$$
|\lambda E-A| = \begin{vmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{vmatrix}
$$

$$
=(\lambda-a_{11})(\lambda-a_{22})\cdots(\lambda-a_{nn})+\cdots.
$$

在上面的行列式的完全展开式中,除主对角线上元素的连乘积一项外,其它 $n!-1$ 项均用省略号代替.因为在$|\lambda E-A|$中只有主对角线上的元素是 λ 的一次多项式,其它全是常数,在用省略号代表的 $n!-1$ 项中的任一项都至少包含一个非主对角线元素:$-a_{ij}$ $(i\neq j)$.此时该项不能包含因子 $\lambda-a_{ii}$和 $\lambda-a_{jj}$(因为行列式的完全展开式中每一项是由取自不同行和不同列的 n 个元素相乘得到的,该项取了 i 行 j 列元素$-a_{ij}$,就再不能取 i 行的其它元素,也不能取 j 列的其它元素).所以,用省略号代表的 $n!-1$ 个项中每一项顶多包含 $n-2$ 个主对角线元素,因而顶多是 λ 的 $n-2$ 次多项式.这样,$f(\lambda)=|\lambda E-A|$中 λ^n 和 λ^{n-1}的系数由主对角线元素连乘

积的展开式得出. 于是,我们有

$$f(\lambda)=\lambda^n-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}+\cdots.$$

这说明:

(1) $f(\lambda)$是 λ 的首项系数为 1 的 n 次多项式.

(2) $f(\lambda)$的 λ^{n-1}项的系数为

$$-(a_{11}+a_{22}+\cdots+a_{nn}).$$

括号中的数恰为矩阵 A 的主对角线元素之和,我们称之为矩阵 A 的**迹**(trace),记作

$$\mathrm{Tr}(A)=a_{11}+a_{22}+\cdots+a_{nn}.$$

(3) $f(\lambda)$的常数项为 $f(0)$,而

$$f(0)=|-A|=(-1)^n|A|.$$

我们知道,一个 n 次多项式在复数域内恰有 n 个根(其中可能有相同的). 设 $f(\lambda)$在复数域内的 n 个根是 $\lambda_1,\lambda_2,\cdots,\lambda_n$,那么,根据根与系数的关系,有

$$\lambda_1+\lambda_2+\cdots+\lambda_n=\mathrm{Tr}(A);$$

$$\lambda_1\cdot\lambda_2\cdot\cdots\cdot\lambda_n=|A|.$$

即矩阵 A 的全体特征根(重根计算在内)之和等于它的迹 $\mathrm{Tr}(A)$,而全体特征根的乘积等于它的行列式$|A|$. 上面阐述的几条简单性质对讨论许多问题都很有用.

具有对角形矩阵的线性变换

设 $\boldsymbol{A}$ 是 n 维线性空间 V 内的一个线性变换. 如果在 V 内存在一组基 $\eta_1,\eta_2,\cdots,\eta_n$,使 $\boldsymbol{A}$ 在这组基下的矩阵成对角形,我们就说 $\boldsymbol{A}$ 的矩阵**可对角化**. 现在我们来考察一下,什么样的线性变换其矩阵可对角化.

定理 1 n 维线性空间 V 内一个线性变换 $\boldsymbol{A}$ 的矩阵可对角化的充分必要条件是,$\boldsymbol{A}$ 有 n 个线性无关的特征向量.

证 (i) 必要性. 若 $\boldsymbol{A}$ 在基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的矩阵成对角形

$$(A\eta_1,A\eta_2,\cdots,A\eta_n)=(\eta_1,\eta_2,\cdots,\eta_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},$$

则有 $$A\eta_1=\lambda_1\eta_1,\quad A\eta_2=\lambda_2\eta_2,\quad \cdots,\quad A\eta_n=\lambda_n\eta_n.$$

于是 $\eta_1,\eta_2,\cdots,\eta_n$ 是 A 的 n 个线性无关特征向量.

(ii) 充分性. 如果 A 有 n 个线性无关特征向量 $\eta_1,\eta_2,\cdots,\eta_n$，把它们取作 V 的一组基，显然，A 在这在组基下的矩阵成对角形. ▌

究竟什么样的线性变换才具有 n 个线性无关的特征向量呢？我们现在给出一个充分条件.

命题 3.2 线性变换 A 的属于不同特征值的特征向量线性无关.

证 取 A 的 k 个不同特征值 $\lambda_1,\lambda_2,\cdots,\lambda_k$，它们分别对应于特征向量 $\xi_1,\xi_2,\cdots,\xi_k$. 对 k 作数学归纳法.

$k=1$ 时，$\xi_1\neq 0$，当然线性无关. 设命题在 $k-1$ 个不同特征值的情况下已经成立，证明 k 个不同特征值的情况下命题也成立.

考察向量等式

$$l_1\xi_1+l_2\xi_2+\cdots+l_k\xi_k=0. \tag{3}$$

两边用 A 作用，利用 A 的线性性质，得

$$l_1A\xi_1+l_2A\xi_2+\cdots+l_kA\xi_k=0.$$

因为 $A\xi_i=\lambda_i\xi_i$，故有

$$l_1\lambda_1\xi_1+l_2\lambda_2\xi_2+\cdots+l_k\lambda_k\xi_k=0. \tag{4}$$

以 λ_1 乘(3)式再与(4)式相减，得

$$l_2(\lambda_1-\lambda_2)\xi_2+\cdots l_k(\lambda_1-\lambda_k)\xi_k=0.$$

按归纳假设，$\xi_2,\cdots,\xi_k$ 线性无关，故

$$l_2(\lambda_1-\lambda_2)=\cdots=l_k(\lambda_1-\lambda_2)=0.$$

因为 k 个特征值互不相同，由上式即得

$$l_2=\cdots=l_k=0.$$

代入(3)式，因 $\xi_1\neq 0$，就有 $l_1=0$. 这就证明 $\xi_1,\xi_2,\cdots,\xi_k$ 是线性无

关的. ∎

推论 如果线性变换 $\boldsymbol{A}$ 有 n 个不同的特征值,**那么**它的矩阵可对角化.

这个推论请读者自行证明.

现在考察一个 n 阶方阵 A. 把它看作复数域 $\mathscr{C}$ 上 n 维向量空间 $\mathscr{C}^n$ 上的一个线性变换

$$\boldsymbol{A}X=AX,\quad X\in\mathscr{C}^n$$

($\mathscr{C}^n$ 中的向量此处表为列形式). 在 $\mathscr{C}^n$ 中取单位向量

$$\varepsilon_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},\quad \varepsilon_2=\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\quad \cdots,\quad \varepsilon_n=\begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}$$

作为一组基. 不难验证,$\boldsymbol{A}$ 在这组基下的矩阵即为 A. 现设 A 的特征多项式

$$f(\lambda)=|\lambda E-A|$$

在复数域内有 n 个不同的根 $\lambda_1,\lambda_2,\cdots,\lambda_n$. 那么,这 n 个复数都是 $\boldsymbol{A}$ 的特征值. 假定 λ_i 对应于特征向量

$$X_i=\begin{pmatrix}x_{1i}\\x_{2i}\\\vdots\\x_{ni}\end{pmatrix}\quad(i=1,2,\cdots,n).$$

由命题 3.2,$X_1,X_2,\cdots,X_n$ 线性无关,所以它们构成 $\mathscr{C}^n$ 的一组基. 由于 $\boldsymbol{A}X_i=\lambda_iX_i(i=1,2,\cdots,n)$,所以 $\boldsymbol{A}$ 在 $X_1,X_2,\cdots,X_n$ 这组基下的矩阵为

$$\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{pmatrix}.$$

而由 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 到 $X_1,X_2,\cdots,X_n$ 的过渡矩阵显然是

$$T=\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{pmatrix},$$

根据命题 2.4，即知

$$T^{-1}AT=\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

所以，如果一个 n 阶方阵 A 的特征多项式 $f(\lambda)$ 有 n 个不同的复根时，A 在复数域内相似于一个对角矩阵.

上面的讨论也适用于矩阵 A 的互不相同特征值少于 n 个，但却有 n 个线性无关的特征向量的情形. 在这种情况下，为了求出可逆矩阵 T，使 $T^{-1}AT$ 成对角形，只要把 A 的 n 个线性无关特征向量找出来，以它们为列向量所排成的 n 阶方阵就是所要寻找的矩阵 T 了. 这时要注意：对角矩阵中特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 的排列次序要与 T 中特征向量（作为 T 的列向量）的排列次序相一致.

习 题 三

1. 设 $\boldsymbol{A}$ 是线性空间 V 内的线性变换，若 $\boldsymbol{A}\alpha=\lambda_0\alpha$，又设 $f(\lambda)=a_0\lambda^m+a_1\lambda^{m-1}+\cdots+a_m$. 证明：

$$f(\boldsymbol{A})\alpha=f(\lambda_0)\alpha.$$

2. 设 $\boldsymbol{A},\boldsymbol{B}$ 是线性空间 V 内的两个线性变换，且 $\boldsymbol{AB}=\boldsymbol{BA}$. 证明：若 $\boldsymbol{A}\alpha=\lambda_0\alpha$，则 $\boldsymbol{B}\alpha\in V_{\lambda_0}$.

3. 设 $\boldsymbol{A}$ 是 n 维线性空间 V 内的一个线性变换，在基 $\varepsilon_1,\varepsilon_2,\cdots\varepsilon_n$ 下的矩阵为

$$A=\begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}.$$

证明：$\boldsymbol{A}$ 只有唯一的特征值 $\lambda_0=0$，且 $V_{\lambda_0}=L(\varepsilon_1)$.

4. 设 **A** 是 n 维线性空间 V 内的一个线性变换. 如果在 V 内存在一组基

$$\varepsilon_1, \cdots, \varepsilon_r, \varepsilon_{r+1}, \cdots, \varepsilon_n,$$

使 **A** 在这组基下的矩阵为如下准对角形

$$A=\begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix},$$

其中 J_1, J_2 分别为 r 阶与 $n-r$ 阶方阵，且

$$J_1=\begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}, \quad J_2=\begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}.$$

证明：**A** 只有一个特征值 $\lambda_0=0$，且 $V_{\lambda_0}=L(\varepsilon_1, \varepsilon_{r+1})$.

5. 设 **A** 是复数域上线性空间 V 内的一个线性变换，且它在某一组基 ε_i 下的矩阵为 A，求 **A** 的全部特征值和每个特征值 λ_i 所属特征子空间 V_{λ_i} 的一组基，其中：

(1) $A=\begin{pmatrix} 3 & 4 \\ 5 & 2 \end{pmatrix}$;　　(2) $A=\begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$;

(3) $A=\begin{pmatrix} 5 & 6 & -3 \\ -1 & 0 & 1 \\ 1 & 2 & -1 \end{pmatrix}$;　　(4) $A=\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$;

(5) $A=\begin{pmatrix} 0 & 2 & 1 \\ -2 & 0 & 3 \\ -1 & -3 & 0 \end{pmatrix}$;　　(6) $A=\begin{pmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & -8 & -2 \end{pmatrix}$;

(7) $A=\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$.

6. 上一题中哪些变换的矩阵可对角化？在可对角化的情况

下，写出基变换的过渡矩阵 T，并验算 $T^{-1}AT$ 为对角形.

7. 设 $\boldsymbol{A}$ 是线性空间 V 内的一个线性变换，存在一个正整数 k，使 $\boldsymbol{A}^k=\boldsymbol{0}$. 证明：$\boldsymbol{A}$ 只有唯一的特征值 $\lambda_0=0$.

8. 设 λ_1,λ_2 是线性变换 $\boldsymbol{A}$ 的两个不同特征值，ξ_1,ξ_2 是分别属于 λ_1,λ_2 的特征向量. 证明：$\xi_1+\xi_2$ 不是 $\boldsymbol{A}$ 的特征向量.

9. 证明：如果线性空间 V 的线性变换 $\boldsymbol{A}$ 以 V 的每个非零向量作为特征向量，则 $\boldsymbol{A}$ 是数乘变换.

10. 设 $\boldsymbol{A}$ 是线性空间 V 内的可逆线性变换.

（1）证明：$\boldsymbol{A}$ 的特征值都不为零；

（2）证明：若 λ 是 $\boldsymbol{A}$ 的一个特征值，则 $1/\lambda$ 是 $\boldsymbol{A}^{-1}$ 的一个特征值.

11. 设 $\boldsymbol{A}$ 是 n 维线性空间 V 内的一个线性变换，在 V 的某一组基下其矩阵成对角形. 又设 $\boldsymbol{A}$ 的全部不同特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_s$. 证明：

$$V=V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_s}.$$

§4 若当标准形

在这一节里，我们介绍复数域上 n 维线性空间 V 内一个线性变换 $\boldsymbol{A}$ 的矩阵的最简单形式. 本节的结果对于非复数域上的线性空间一般是不适用的.

若当标准形介绍

我们在前面已经说过，并不是对任一线性变换 $\boldsymbol{A}$，都可以在线性空间 V 内找出一组基，使 $\boldsymbol{A}$ 在这一组基下的矩阵成对角形. 但是我们可以证明，对于复数域上的线性空间 V，一定可以找出一组基，使 $\boldsymbol{A}$ 在这组基下的矩阵成为准对角形，而且在主对角线上的小块方阵具有很规则的形状.

定义 形如

$$J=\begin{pmatrix}\lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda\end{pmatrix}$$

的方阵称为**若当**(Jordan)**块**,其中 λ 是一个复数. 由若干个若当块组成的准对角矩阵

$$J=\begin{pmatrix}J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s\end{pmatrix},\quad J_i=\begin{pmatrix}\lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i\end{pmatrix}\quad (i=1,2,\cdots,s)$$

称为**若当形矩阵**.

例如

$$\begin{pmatrix}2 & 1 & 0\\ 0 & 2 & 1\\ 0 & 0 & 2\end{pmatrix},\quad \begin{pmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\end{pmatrix},\quad \begin{pmatrix}\mathrm{i} & 1\\ 0 & \mathrm{i}\end{pmatrix}$$

都是若当块. 特别地,一阶方阵都是若当块. 而下面的矩阵

$$\begin{pmatrix}1 & 1 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 4 & 0 & 0 & 0\\ 0 & 0 & 0 & -\mathrm{i} & 1 & 0\\ 0 & 0 & 0 & 0 & -\mathrm{i} & 1\\ 0 & 0 & 0 & 0 & 0 & -\mathrm{i}\end{pmatrix}$$

是一个若当形矩阵. 特别地,所有对角矩阵都是若当形矩阵,其主对角线上的若当块都是一阶的方阵.

定理 2 设 V 是复数域上的一个 n 维线性空间,$\boldsymbol{A}$ 是 V 内一个线性变换. 则在 V 中可以找到一组基,使 $\boldsymbol{A}$ 在这组基下的矩阵是若当形矩阵,而且其主对角线上的若当块除去排列的次序可以不同外,是被 $\boldsymbol{A}$ 唯一决定的.

$\boldsymbol{A}$ 的若当形矩阵称为 $\boldsymbol{A}$ 的**若当标准形**.

这个定理的证明比较复杂，本书从略. 下面对这个定理的意思作一些解释.

首先考察 $\boldsymbol{A}$ 在某一组基

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$$

下的矩阵恰好是一个若当块

$$J=\begin{pmatrix} \lambda_0 & 1 & & \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0 \end{pmatrix}$$

的特殊情形. 这时有

$$\begin{cases} \boldsymbol{A}\varepsilon_1=\lambda_0\varepsilon_1, \\ \boldsymbol{A}\varepsilon_2=\varepsilon_1+\lambda_0\varepsilon_2, \\ \boldsymbol{A}\varepsilon_3=\varepsilon_2+\lambda_0\varepsilon_3, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \boldsymbol{A}\varepsilon_n=\varepsilon_{n-1}+\lambda_0\varepsilon_n. \end{cases}$$

第一个等式说明 λ_0 是 $\boldsymbol{A}$ 的特征值，而 ε_1 则是对应的特征向量. 因为 $\boldsymbol{A}$ 的特征多项式为

$$f(\lambda)=|\lambda E-J|=\begin{vmatrix} \lambda-\lambda_0 & -1 & & \\ & \lambda-\lambda_0 & \ddots & \\ & & \ddots & -1 \\ & & & \lambda-\lambda_0 \end{vmatrix}$$

$$=(\lambda-\lambda_0)^n,$$

故 $\boldsymbol{A}$ 只有唯一的特征值 λ_0. 不难证明：$\boldsymbol{A}$ 的特征子空间 $V_{\lambda_0}=L(\varepsilon_1)$. 这留给读者作为练习.

如果令 $\boldsymbol{B}=\boldsymbol{A}-\lambda_0\boldsymbol{E}$，那么 $\boldsymbol{B}$ 在 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 这组基下的矩阵为

$$\begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}.$$

这表示

$$\boldsymbol{B}\varepsilon_1=0,\ \boldsymbol{B}\varepsilon_2=\varepsilon_1,\ \boldsymbol{B}\varepsilon_3=\varepsilon_2,\ \cdots,\ \boldsymbol{B}\varepsilon_n=\varepsilon_{n-1}.$$

即在 $\boldsymbol{B}$ 的作用下,基向量的下角标依次减 1,第 1 个基向量 ε_1 则变为零.

在一般情况下,设 $\boldsymbol{A}$ 在 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 这组基下的矩阵为若当形

$$J=\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{pmatrix},\quad J_i=\begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}_{n_i\times n_i}.$$

这时,基向量可以相应地分成 s 段

$$\varepsilon_{11},\varepsilon_{12},\cdots,\varepsilon_{1n_1},\varepsilon_{21},\varepsilon_{22},\cdots,\varepsilon_{2n_2},\cdots,\varepsilon_{s1},\varepsilon_{s2},\cdots,\varepsilon_{sn_s}.$$

$\boldsymbol{A}$ 在第 i 段 $\varepsilon_{i1},\varepsilon_{i2},\cdots,\varepsilon_{in_i}$ 的作用相当于在子空间 $L(\varepsilon_{i1},\varepsilon_{i2},\cdots,\varepsilon_{in_i})$ 中以 J_i 为矩阵的线性变换的作用,所以,可以把它归结为上面所指出的特殊情况. 由此可知,$\boldsymbol{A}$ 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_s$(其中可能有相同的). 而每个特征子空间 V_{λ_i} 由特征值相同的若当块的第一个基向量 ε_{k1}(对一切满足 $\lambda_k=\lambda_i$ 的 k)生成. 因为 J 是上三角矩阵,不难算出 $\boldsymbol{A}$ 的特征多项式为

$$f(\lambda)=|\lambda E-J|=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}.$$

任意一个 n 阶方阵都可以看作复数域上某个 n 维线性空间 V 中一个线性变换 $\boldsymbol{A}$ 在一组基下的矩阵. 根据定理 2,我们有:

定理 3 任一 n 阶方阵 A 在复数域内都相似于一个若当形矩阵,且除了若当块的排列次序可有不同外,这样的若当形矩阵是被 A 唯一决定的,它称为 A 的**若当标准形**.

*二阶和三阶矩阵的若当形

在某些实际问题中需要具体计算低阶的特别是二、三阶的矩阵的若当标准形,因此,我们在这一段里给出定理 2 在 $n=2,3$ 时的证明. 在证明的过程中同时给出了二、三阶矩阵的若当形的具体计算方法.

我们首先证明一个比定理 2 稍弱一点的命题. 这个命题对我们研究低阶矩阵的若当形的计算方法是有用的.

命题 4.1 设 $\boldsymbol{A}$ 是数域 K 上的 n 维线性空间 V 内的一个线

性变换，如果 $\boldsymbol{A}$ 的特征多项式 $f(\lambda)$ 的根全在 K 内，那么，$\boldsymbol{A}$ 的任意 r 个线性无关特征向量 $\eta_1,\eta_2,\cdots,\eta_r(r\geqslant 1)$ 都可扩充成 V 的一组基

$$\eta_1,\cdots,\eta_r,\eta_{r+1},\cdots,\eta_n,$$

使 $\boldsymbol{A}$ 在这组基下的矩阵为如下分块形式

$$A=\begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix},$$

其中 A_1,A_3 分别为 r 阶与 $n-r$ 阶方阵，且

$$A_1=\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_r \end{pmatrix},\quad A_3=\begin{pmatrix} a_{r+1\,r+1} & \cdots & a_{r+1\,n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{pmatrix}$$

（当 $r=n$ 时，$A=A_1$ 为对角矩阵）.

证 对 V 的维数 n 作数学归纳法.

$n=1$ 时命题是显然的. 设对一切维数 $\leqslant n-1$ 的线性空间命题均成立，证明对 n 维线性空间命题也成立.

现设 $\eta_1,\cdots,\eta_r$ 是 $\boldsymbol{A}$ 的 r 个线性无关特征向量，不妨设 $1\leqslant r<n$（若 $r=n$，命题已成立）. 把它们扩充成 V 的一组基

$$\eta_1,\cdots,\eta_r,\alpha_{r+1},\cdots,\alpha_n.$$

在这组基下 $\boldsymbol{A}$ 的矩阵可写成

$$A_1=\begin{pmatrix} \lambda_1 & & 0 & b_{1\,r+1} & \cdots & b_{1n} \\ & \ddots & & \vdots & & \vdots \\ 0 & & \lambda_r & b_{r\,r+1} & \cdots & b_{rn} \\ & & & b_{r+1\,r+1} & \cdots & b_{r+1\,n} \\ & 0 & & \vdots & & \vdots \\ & & & b_{n\,r+1} & \cdots & b_{nn} \end{pmatrix}.$$

命

$$B=\begin{pmatrix} b_{r+1\,r+1} & \cdots & b_{r+1\,n} \\ \vdots & & \vdots \\ b_{n\,r+1} & \cdots & b_{nn} \end{pmatrix}.$$

$\boldsymbol{A}$ 的特征多项式 $f(\lambda)=|\lambda E-A_1|=(\lambda-\lambda_1)\cdots(\lambda-\lambda_r)|\lambda E-B|$. $f(\lambda)$ 的根全在 K 内，故 $|\lambda E-B|$ 的根也全在 K 内. 令

$$M=L(\eta_1,\cdots,\eta_r);\quad N=L(\alpha_{r+1},\cdots,\alpha_n).$$

在子空间 N 内定义线性变换 $\boldsymbol{B}$ 如下：

$$\boldsymbol{B}\alpha_{r+1}=b_{r+1\,r+1}\alpha_{r+1}+b_{r+2\,r+1}\alpha_{r+2}+\cdots+b_{n\,r+1}\alpha_n,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\boldsymbol{B}\alpha_n=b_{r+1\,n}\alpha_{r+1}+b_{r+2\,n}\alpha_{r+2}+\cdots+b_{nn}\alpha_n.$$

$\boldsymbol{B}$ 在基 $\alpha_{r+1},\cdots,\alpha_n$ 下的矩阵为 $\boldsymbol{B}$，其特征多项式为 $|\lambda E-B|$，根全在 K 内. 而 $\dim N=n-r<n$，按归纳假设，在 N 内存在一组基 $\eta_{r+1},\cdots,\eta_n$，使 $\boldsymbol{B}$ 在这组基下的矩阵成上三角形

$$A_3=\begin{pmatrix} a_{r+1\,r+1} & \cdots & a_{r+1\,n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{pmatrix}.$$

另一方面，因为

$$\boldsymbol{A}\alpha_{r+k}=\boldsymbol{B}\alpha_{r+k}+\beta_{r+k}\quad(k=1,2,\cdots,n-r),$$

其中

$$\beta_{r+k}=b_{1\,r+k}\eta_1+\cdots+b_{r\,r+k}\eta_r\in M,$$

所以对 N 内任一向量 $\alpha=c_{r+1}\alpha_{r+1}+\cdots+c_n\alpha_n$，有

$$\begin{aligned}\boldsymbol{A}\alpha &=c_{r+1}\boldsymbol{A}\alpha_{r+1}+\cdots+c_n\boldsymbol{A}\alpha_n\\ &=c_{r+1}(\boldsymbol{B}\alpha_{r+1}+\beta_{r+1})+\cdots+c_n(\boldsymbol{B}\alpha_n+\beta_n)\\ &=(c_{r+1}\boldsymbol{B}\alpha_{r+1}\cdots+c_n\boldsymbol{B}\alpha_n)+(c_{r+1}\beta_{r+1}+\cdots+c_n\beta_n)\\ &=\boldsymbol{B}\alpha+\beta\quad(\beta\in M).\end{aligned}$$

特别地

$$\boldsymbol{A}\eta_{r+k}=\boldsymbol{B}\eta_{r+k}+\gamma_{r+k}\quad(\gamma_{r+k}\in M).$$

因为向量组 $\eta_1,\cdots,\eta_r,\eta_{r+1},\cdots,\eta_n$ 与 $\eta_1,\cdots,\eta_r,\alpha_{r+1},\cdots,\alpha_n$ 等价，故它也是 V 的一组基. $\boldsymbol{A}$ 在这组基下的矩阵即具有命题中所要求的形式. ▍

在上面这个命题中，矩阵 A_3 的主对角线上的元素全都是 $\boldsymbol{A}$ 的特征多项式 $f(\lambda)$ 的根，所以全是 $\boldsymbol{A}$ 的特征值.

对于复数域 $\mathscr{C}$ 上的 n 维线性空间来说，命题 4.1 的条件自然具备，因而它总是可以应用的. 根据这个命题，只要设法先把一个线性变换 $\boldsymbol{A}$ 的全部线性无关特征向量找出来，就可以把它的矩阵

大大简化，这就为我们寻求它的若当标准形提供了方便的条件. 特别是对低维的线性空间（或低阶的矩阵）是如此.

一、二阶矩阵的若当标准形

命题 4.2 设 $\boldsymbol{A}$ 是复数域上二维线性空间 V 内的一个线性变换，则在 V 内存在一组基 η_1,η_2，使 $\boldsymbol{A}$ 在这组基下的矩阵成下列两种形式之一：

$$\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix},\quad \begin{pmatrix}\lambda_1 & 1\\ 0 & \lambda_1\end{pmatrix}.$$

证 若 $\boldsymbol{A}$ 有两个线性无关的特征向量，则由定理 1 知 $\boldsymbol{A}$ 的矩阵可对角化. 下面设 $\boldsymbol{A}$ 仅有一个线性无关特征向量. 由命题 3.2 的推论知，$\boldsymbol{A}$ 的特征多项式仅有一个重根 λ_1. 设 λ_1 对应的特征向量为 η_1'，把它扩充成 V 的一组基 η_1',η_2. 在此组基下 $\boldsymbol{A}$ 的矩阵应为

$$\begin{pmatrix}\lambda_1 & a\\ 0 & \lambda_1\end{pmatrix}.$$

因为 $\boldsymbol{A}\eta_2=\lambda_1\eta_2+a\eta_1'$，而 η_2 不是 $\boldsymbol{A}$ 的特征向量，故 $a\neq 0$. 命 $\eta_1=a\eta_1'$，则 η_1,η_2 是 V 的一组基，且

$$\boldsymbol{A}\eta_1=\lambda_1\eta_1,\quad \boldsymbol{A}\eta_2=\lambda_1\eta_2+\eta_1.$$

故 $\boldsymbol{A}$ 在基 η_1,η_2 下的矩阵为

$$\begin{pmatrix}\lambda_1 & 1\\ 0 & \lambda_1\end{pmatrix}.\ \blacksquare$$

在这里，$\boldsymbol{A}$ 的若当形的唯一性（除了差一个若当块的排列次序之外）是显然的. 命题的证明过程同时就给出了基 η_1,η_2 的寻求方法.

任给一个二阶矩阵

$$A=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix},$$

把它看作复数域上某个二维线性空间 V 内一个线性变换 $\boldsymbol{A}$ 在基

$\varepsilon_1,\varepsilon_2$ 下的矩阵. 按命题 4.2 的办法找出的新基 η_1,η_2,使 $\boldsymbol{A}$ 在新基下的矩阵成若当形. 再求出两组基之间的过渡矩阵 T:

$$(\eta_1,\eta_2)=(\varepsilon_1,\varepsilon_2)T,$$

则 $T^{-1}AT$ 即为 A 的若当标准形.

二、三阶矩阵的若当标准形

命题 4.3 设 $\boldsymbol{A}$ 是复数域上三维线性空间 V 内的一个线性变换,则在 V 内存在一组基 η_1,η_2,η_3,使 $\boldsymbol{A}$ 在这组基下的矩阵成下列三种形式之一:

$$\begin{pmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3\end{pmatrix},\quad \begin{pmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 1\\ 0 & 0 & \lambda_2\end{pmatrix},\quad \begin{pmatrix}\lambda_1 & 1 & 0\\ 0 & \lambda_1 & 1\\ 0 & 0 & \lambda_1\end{pmatrix}.$$

证 如果 $\boldsymbol{A}$ 有三个线性无关的特征向量 η_1,η_2,η_3,那么 $\boldsymbol{A}$ 在这组基下的矩阵即为对角形. 下面假设不是这种情况. 我们分两种情况讨论.

(1) 如果 $\boldsymbol{A}$ 恰有两个线性无关特征向量 η_1',η_2',且

$$\boldsymbol{A}\eta_1'=\lambda_1\eta_1',\quad \boldsymbol{A}\eta_2'=\lambda_2\eta_2',$$

此时 $\boldsymbol{A}$ 最多只能有两个不同的特征值. 不妨设 $\boldsymbol{A}$ 的特征多项式 $f(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)^2$. 添加 η_3',使 η_1',η_2',η_3' 成为 V 的一组基. 又设

$$\boldsymbol{A}\eta_3'=a\eta_1'+b\eta_2'+\lambda_2\eta_3'.$$

因为 η_3' 不是 $\boldsymbol{A}$ 的特征向量,故 a,b 不全为零.

(i) 若 $\lambda_1=\lambda_2$,则选取 a',b',使

$$\begin{vmatrix}a & b\\ a' & b'\end{vmatrix}\neq 0.$$

令

$$\eta_1=a\eta_1'=a'\eta_1'+b'\eta_2',\quad \eta_2=a\eta_1'+b\eta_2',\quad \eta_3=\eta_3'.$$

不难验证 η_1,η_2,η_3 是 V 的一组基,且

$$\boldsymbol{A}\eta_1=\lambda_1\eta_1,\quad \boldsymbol{A}\eta_2=\lambda_2\eta_2,\quad \boldsymbol{A}\eta_3=\eta_2+\lambda_2\eta_3.$$

故 $\boldsymbol{A}$ 在这组基下的矩阵为

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{pmatrix} \quad (\lambda_1 = \lambda_2).$$

(ii) 若 $\lambda_1 \neq \lambda_2$，此时如果 $b=0$，则 $\boldsymbol{A}\eta'_3 = a\eta'_1 + \lambda_2\eta'_3$. 于是有

$$\boldsymbol{A}[a\eta'_1 + (\lambda_2 - \lambda_1)\eta'_3] = \lambda_2[a\eta'_1 + (\lambda_2 - \lambda_1)\eta'_3],$$

这说明 $\boldsymbol{A}$ 有 $\eta'_1, \eta'_2, a\eta'_1 + (\lambda_2 - \lambda_1)\eta'_3$ 三个线性无关的特征向量，与假设矛盾. 故 $b \neq 0$. 令

$$\eta_1 = \eta'_1 \quad \eta_2 = b\eta'_2, \quad \eta_3 = \eta'_3 - \frac{a}{\lambda_1 - \lambda_2}\eta'_1.$$

我们有

$$\boldsymbol{A}\eta_1 = \lambda_1\eta_1, \quad \boldsymbol{A}\eta_2 = = \lambda_2\eta_2,$$

$$\begin{aligned} \boldsymbol{A}\eta_3 &= \boldsymbol{A}\eta'_3 - \frac{a}{\lambda_1 - \lambda_2}\boldsymbol{A}\eta'_1 = a\eta'_1 + b\eta'_2 + \lambda_2\eta'_3 - \frac{a}{\lambda_1 - \lambda_2}\lambda_1\eta'_1 \\ &= \eta_2 + \lambda_2\eta_3. \end{aligned}$$

故 $\boldsymbol{A}$ 在 η_1, η_2, η_3 这组基下的矩阵为

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{pmatrix}.$$

(2) 如果 $\boldsymbol{A}$ 仅有一个线性无关的特征向量 η'_1，此时 $\boldsymbol{A}$ 仅有一个三重特征值 λ_1. 由命题 4.1，可以把 η'_1 扩充成 $\boldsymbol{V}$ 的一组基 η'_1，η'_2, η'_3，使

$$\begin{aligned} \boldsymbol{A}\eta'_1 &= \lambda_1\eta'_1 \\ \boldsymbol{A}\eta'_2 &= a\eta'_1 + \lambda_1\eta'_2, \\ \boldsymbol{A}\eta'_3 &= b\eta'_1 + c\eta'_2 + \lambda_1\eta'_3. \end{aligned}$$

若 $c=0$，则找一组不全为零的数 k, l，使 $ak+bl=0$. 于是

$$\begin{aligned} \boldsymbol{A}(k\eta'_2 + l\eta'_3) &= k(a\eta'_1 + \lambda_1\eta'_2) + l(b\eta'_1 + \lambda_1\eta'_3) \\ &= (ak+bl)\eta'_1 + \lambda_1(k\eta'_2 + l\eta'_3) = \lambda_1(k\eta'_2 + l\eta'_3). \end{aligned}$$

即 $\boldsymbol{A}$ 有两个线性无关特征向量 $\eta'_1, k\eta'_2 + l\eta'_3$，与假设矛盾. 故必有 $c \neq 0$. 令

$$\eta_2=b\eta_1'+c\eta_2', \quad \eta_3=\eta_3',$$

则

$$A\eta_3=\eta_2+\lambda_1\eta_3,$$
$$A\eta_2=ac\eta_1'+\lambda_1\eta_2.$$

因 η_2 不是 A 的特征向量，故 $ac\neq 0$. 令 $\eta_1=ac\eta_1'$，则在基 η_1,η_2,η_3 下的 A 的矩阵为

$$\begin{pmatrix}\lambda_1 & 1 & 0\\ 0 & \lambda_1 & 1\\ 0 & 0 & \lambda_1\end{pmatrix}. \blacksquare$$

上面命题的证明过程已经指出了基向量 η_1,η_2,η_3 的寻求方法，我们对此再作一点说明：在上述证明过程的(2)中，找出了 A 的一个线性无关特征向量 η_1' 后，还要求出所需要的 η_1,η_2,η_3，可采用如下办法：先补充 α_2,α_3，使 $\eta_1',\alpha_2,\alpha_3$ 构成 V 的一组基，且设 $B=A-\lambda_1E$. 因为 B 在 η_1,η_2,η_3 这组基下的矩阵为

$$\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{pmatrix},$$

故知 $B^2\neq \mathbf{0}$. 而 $B^2\eta_1'=0$，故 $B^2\alpha_2$ 与 $B^2\alpha_3$ 中必有一个不为零. 设 $B^2\alpha_2\neq 0$，于是令

$$\eta_3=\alpha_2, \quad \eta_2=B\alpha_2, \quad \eta_1=B^2\alpha_2,$$

则 η_1,η_2,η_3 组成 V 的一组基，且 B 在这组基下的矩阵为若当形，从而 A 在这组基下的矩阵也是若当形(参看本章习题二的第 6,7 两题).

从命题的证明可以看出，A 的三种若当形分别对应于 A 有三个、二个、一个线性无关特征向量的情况，由此可知若当形是唯一的.

如果给定一个三阶方阵 A，把它看作复数域上某个三维线性空间 V 内线性变换 A 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵，按上述办法找出新

基 η_1,η_2,η_3，使 $\boldsymbol{A}$ 在新基下的矩阵成若当形. 再求两组基之间的过渡矩阵 T

$$(\eta_1,\eta_2,\eta_3)=(\varepsilon_1,\varepsilon_2,\varepsilon_3)T,$$

则 $T^{-1}AT$ 即为 A 的若当形矩阵.

若当标准形的一个应用

矩阵的若当标准形是一个有用的工具，我们现在介绍它对一阶线性常系数微分方程组求解问题的应用.

设 $y_1(x),y_2(x),\cdots,y_n(x)$ 是 n 个未知函数，满足如下一阶线性常微分方程组

$$\begin{cases}\dfrac{\mathrm{d}y_1}{\mathrm{d}x}=a_{11}y_1+a_{12}y_2+\cdots+a_{1n}y_n,\\ \dfrac{\mathrm{d}y_2}{\mathrm{d}x}=a_{21}y_1+a_{22}y_2+\cdots+a_{2n}y_n,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \dfrac{\mathrm{d}y_n}{\mathrm{d}x}=a_{n1}y_1+a_{n2}y_2+\cdots+a_{nn}y_n,\end{cases} \tag{1}$$

其中 a_{ij} 为复常数. 命

$$Y=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix},\quad \frac{\mathrm{d}Y}{\mathrm{d}x}=\frac{\mathrm{d}}{\mathrm{d}x}\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}=\begin{pmatrix}y_1'\\y_2'\\\vdots\\y_n'\end{pmatrix},$$

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}.$$

方程组(1)可以写成

$$\frac{\mathrm{d}Y}{\mathrm{d}x}=AY. \tag{2}$$

根据定理 3,存在一个复的可逆矩阵 T,使

$$T^{-1}AT=J=\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{pmatrix},\quad J_i=\begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}.$$

做线性变数替换

$$Z=\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}=T^{-1}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}=T^{-1}Y.$$

于是

$$Y=TZ,\ \frac{\mathrm{d}Y}{\mathrm{d}x}=T\frac{\mathrm{d}Z}{\mathrm{d}x}.$$

代入(2)式,得

$$T\frac{\mathrm{d}Z}{\mathrm{d}x}=ATZ,$$

则有

$$\frac{\mathrm{d}Z}{\mathrm{d}x}=T^{-1}ATZ=JZ. \tag{3}$$

这就把方程组(1)大大简化了.(3)式的通解很容易求出.求出 Z 以后,再从 $Y=TZ$ 即可得到原方程组(1)的通解.

为了帮助读者了解以若当形矩阵 J 为系数矩阵的微分方程组(3)的求解方法,我们看一个具体例子:设 $n=4$,

$$J=\begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 1 & 0 \\ 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}.$$

方程组(3)现在是

$$\begin{pmatrix} z_1' \\ z_2' \\ z_3' \\ z_4' \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 1 & 0 \\ 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix}.$$

具体写出，就是

$$\begin{cases} \dfrac{\mathrm{d}z_1}{\mathrm{d}x} = \lambda_1 z_1, & (4) \\ \dfrac{\mathrm{d}z_2}{\mathrm{d}x} = \qquad \lambda_2 z_2 + z_3, & (5) \\ \dfrac{\mathrm{d}z_3}{\mathrm{d}x} = \qquad\qquad \lambda_2 z_3 + z_4, & (6) \\ \dfrac{\mathrm{d}z_4}{\mathrm{d}x} = \qquad\qquad\qquad \lambda_2 z_4. & (7) \end{cases}$$

从方程(4)可立即解出 z_1；从(7)可立即解出 z_4，再代入(6)解出 z_3，最后代入(5)求出 z_2. 方程(5)，(6)，(7)对应于若当形矩阵 J 中的一个若当块. 显然，对任一个若当块都可以用这种办法处理.

习 题 四

1. 设线性空间 V 内的线性变换 $\boldsymbol{A}$ 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r, \varepsilon_{r+1}, \cdots, \varepsilon_n$ 下的矩阵为如下若当形

$$A = \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix},$$

其中

$$J_1 = \begin{pmatrix} \lambda_1 & 1 & & \\ & \lambda_1 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_1 \end{pmatrix}_{r \times r}, \quad J_2 = \begin{pmatrix} \lambda_2 & 1 & & \\ & \lambda_2 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_2 \end{pmatrix}.$$

证明：

(i) 若 $\lambda_1 = \lambda_2$，则 $V_{\lambda_1} = L(\varepsilon_1, \varepsilon_{r+1})$；

(ii) 若 $\lambda_1 \neq \lambda_2$，则 $V_{\lambda_1} = L(\varepsilon_1), V_{\lambda_2} = L(\varepsilon_2)$.

2. 给定若当块矩阵

$$J=\begin{pmatrix}\lambda_0 & 1 & & \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0\end{pmatrix}_{k\times k},$$

证明：$(J-\lambda_0 E)^k=0$.

3. 给定若当形矩阵

$$J=\begin{pmatrix}J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s\end{pmatrix},\quad J_i=\begin{pmatrix}\lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i\end{pmatrix}_{n_i\times n_i},$$

其特征多项式设为 $f(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}=0$. 证明：

$$f(J)=(J-\lambda_1 E)^{n_1}(J-\lambda_2 E)^{n_2}\cdots(J-\lambda_s E)^{n_s}=0.$$

4. 证明哈密顿-凯莱(Hamilton-Cayley)定理：设 A 是一个 n 阶方阵，$f(\lambda)$ 为 A 的特征多项式，则 $f(A)=0$.

5. 设 J 是若当块矩阵

$$J=\begin{pmatrix}\lambda_0 & 1 & & \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0\end{pmatrix},$$

证明 J 与 J' 相似.

6. 设 J 是一个若当形矩阵，证明：J 与 J' 相似.

7. 设 A 是一个 n 阶方阵，证明：A 与 A' 相似.

8. 求下列矩阵的若当标准形，并求出化 A 为若当形的过渡矩阵 T：

(1) $A=\begin{pmatrix}1 & 2 & 0\\ 0 & 2 & 0\\ -2 & -2 & -1\end{pmatrix}$；　(2) $A=\begin{pmatrix}3 & 0 & 8\\ 3 & -1 & 6\\ -2 & 0 & -5\end{pmatrix}$；

(3) $A=\begin{pmatrix}4 & 5 & -2\\ -2 & -2 & 1\\ -1 & -1 & 1\end{pmatrix}$；　(4) $A=\begin{pmatrix}1 & -1 & 2\\ 3 & -3 & 6\\ 2 & -2 & 4\end{pmatrix}$；

(5) $A=\begin{pmatrix}1 & 1 & -1\\ -3 & -3 & 3\\ -2 & -2 & 2\end{pmatrix}$；　(6) $A=\begin{pmatrix}-4 & 2 & 10\\ -4 & 3 & 7\\ -3 & 1 & 7\end{pmatrix}$.

9. 解下列常微分方程组

(1) $\begin{cases} \dfrac{dy_1}{dx}=5y_1-y_2, \\ \dfrac{dy_2}{dx}=2y_1+y_2. \end{cases}$

(2) $\begin{cases} \dfrac{dy_1}{dx}=2y_1+y_2+y_3, \\ \dfrac{dy_2}{dx}=y_1+2y_2+y_3, \\ \dfrac{dy_3}{dx}=2y_1+2y_2. \end{cases}$

*§5 不变子空间

研究线性变换的一个重要方法,是把它所作用的空间进行分解. 在这一节里,我们介绍与此有关的一些基本概念.

定义 设 $\boldsymbol{A}$ 是线性空间 V 内的一个线性变换. 如果 M 是 V 的一个子空间,且对任意 $\alpha\in M$,有 $\boldsymbol{A}\alpha\in M$,则称 M 为 $\boldsymbol{A}$ 的一个**不变子空间**. 这时 $\boldsymbol{A}$ 可以看作 M 内的一个线性变换,称为 $\boldsymbol{A}$ 在 M 内的限制,记作 $\boldsymbol{A}|_M$.

显然,零子空间 $\{0\}$ 和 V 本身都是 $\boldsymbol{A}$ 的不变子空间,称它们为 $\boldsymbol{A}$ 的**平凡不变子空间**. 除此之外的不变子空间称为 $\boldsymbol{A}$ 的**非平凡不变子空间**. 例如,当 λ 是 $\boldsymbol{A}$ 的一个特征值时,V_λ 是 $\boldsymbol{A}$ 的一个不变子空间,且 $V_\lambda\neq\{0\}$. 如果 $\boldsymbol{A}$ 不是数乘变换,则 $V_\lambda\neq V$,此时 V_λ 就是 $\boldsymbol{A}$ 的一个非平凡不变子空间.

设 M 是 $\boldsymbol{A}$ 的一个非平凡不变子空间,在 M 内取一组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r,$$

扩充成 V 的一组基

$$\varepsilon_1,\cdots,\varepsilon_r,\varepsilon_{r+1},\cdots,\varepsilon_n.$$

按照不变子空间的定义,有

$$A\varepsilon_1=a_{11}\varepsilon_1+a_{21}\varepsilon_2+\cdots a_{r1}\varepsilon_r,$$

$$A\varepsilon_2=a_{12}\varepsilon_1+a_{22}\varepsilon_2+\cdots+a_{r2}\varepsilon_r,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$A\varepsilon_r=a_{1r}\varepsilon_1+a_{2r}\varepsilon_2+\cdots+a_{rr}\varepsilon_r,$$

$$A\varepsilon_{r+1}=a_{1\,r+1}\varepsilon_1+a_{2\,r+1}\varepsilon_2+\cdots+a_{r\,r+1}\varepsilon_r$$
$$+a_{r+1\,r+1}\varepsilon_{r+1}\cdots+a_{n\,r+1}\varepsilon_n,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$A\varepsilon_n=a_{1n}\varepsilon_1+a_{2n}\varepsilon_2+\cdots+a_{rn}\varepsilon_r$$
$$+a_{r+1\,n}\varepsilon_{r+1}+\cdots+a_{nn}\varepsilon_n.$$

故 $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵有如下分块形式

$$A=\begin{pmatrix}A_{11} & A_{12}\\ 0 & A_{22}\end{pmatrix},\qquad A_{11}=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1r}\\ a_{21} & a_{22} & \cdots & a_{2r}\\ \vdots & \vdots & & \vdots\\ a_{r1} & a_{r2} & \cdots & a_{rr}\end{pmatrix}.$$

这就把线性变换 $\boldsymbol{A}$ 的矩阵简化了.这对我们研究线性变换是有利的.

当我们能找到 $\boldsymbol{A}$ 的另一个不变子空间 N,使

$$V=M\oplus N$$

时,只要取 $\varepsilon_{r+1},\cdots,\varepsilon_n$ 为 N 的一组基,则 $\boldsymbol{A}$ 在 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 这组基下的矩阵就成准对角形

$$\begin{pmatrix}A_{11} & 0\\ 0 & A_{22}\end{pmatrix}.$$

事实上,我们有更一般的结果:

命题 5.1 设 $\boldsymbol{A}$ 是线性空间 V 内的一个线性变换.在 V 内存在一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,使 $\boldsymbol{A}$ 在这组基下的矩阵成准对角形的充分必要条件是,V 可以分解为 $\boldsymbol{A}$ 的不变子空间 $M_1,M_2,\cdots,M_s$ 的直和

$$V=M_1\oplus M_2\oplus\cdots\oplus M_s.$$

证 (i) 必要性.若 $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下成上述准对角形,则

把这组基相应分成 s 段

$$\varepsilon_{11},\varepsilon_{12},\cdots,\varepsilon_{1n_1},\varepsilon_{21},\varepsilon_{22},\cdots,\varepsilon_{2n_2},\cdots,\varepsilon_{s1},\varepsilon_{s2},\cdots,\varepsilon_{sn_s},$$

其中 $n_i(i=1,2,\cdots,s)$ 为 A_i 的阶，这时应有

$$(\boldsymbol{A}\varepsilon_{i1},\boldsymbol{A}\varepsilon_{i2},\cdots,\boldsymbol{A}\varepsilon_{in_i})=(\varepsilon_{i1},\varepsilon_{i2},\cdots,\varepsilon_{in_i})A_i.$$

令 $M_i=L(\varepsilon_{i1},\varepsilon_{i2},\cdots,\varepsilon_{in_i})$，则 M_i 为 $\boldsymbol{A}$ 的不变子空间，且

$$\dim M_i=n_i,\quad \sum_{i=1}^{s}M_i=V.$$

而

$$\dim\sum_{i=1}^{s}M_i=\dim V=n=\sum_{i=1}^{s}n_i=\sum_{i=1}^{s}\dim M_i,$$

根据第四章定理 3，有

$$V=M_1\oplus M_2\oplus\cdots\oplus M_s.$$

(ii) 充分性. 设

$$V=M_1\oplus M_2\oplus\cdots\oplus M_s.$$

在每个 M_i 中取一组基，合并后即成 V 的一组基(见第四章习题第 9 题)$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，$\boldsymbol{A}$ 在这组基下的矩阵即为准对角矩阵. ▌

如果 $\boldsymbol{A}$ 在某一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵成若当形，那么，V 就被相应地分解成 $\boldsymbol{A}$ 的不变子空间的直和

$$V=M_1\oplus M_2\oplus\cdots\oplus M_s.$$

而 $\boldsymbol{A}$ 限制在每个 M_i 内，其矩阵在 M_i 的适当的一组基(可由 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 分段得出)下的矩阵是若当块.

最后，我们介绍商空间上的诱导变换的概念. 设 V 是数域 K 上的 n 维线性空间，$\boldsymbol{A}$ 是 V 内的一个线性变换，M 是 $\boldsymbol{A}$ 的一个不变子空间. 考察商空间 V/M，它的元素是 V 内模 M 的一个同余类 $\alpha+M$. 我们在 V/M 内定义一个变换 $\bar{\boldsymbol{A}}$ 如下：

$$\bar{\boldsymbol{A}}(\alpha+M)=\boldsymbol{A}\alpha+M. \tag{1}$$

这样定义变换 $\bar{\boldsymbol{A}}$ 并不会由于 $\alpha+M$ 的代表选取不同而得到不同的结果. 因为：若 $\alpha'\equiv\alpha(\mathrm{mod}M)$，即 $\alpha'=\alpha+\theta(\theta\in M)$，那么，有

$$\boldsymbol{A}\alpha'=\boldsymbol{A}\alpha+\boldsymbol{A}\theta.$$

由于 M 是 $\boldsymbol{A}$ 的不变子空间，故 $\boldsymbol{A}\theta\in M$. 从而 $\boldsymbol{A}\alpha'\equiv\boldsymbol{A}\alpha(\mathrm{mod}M)$，于是 $\boldsymbol{A}\alpha'+M=\boldsymbol{A}\alpha+M$. 这表示

$$\boldsymbol{A}(\alpha'+M)=\boldsymbol{A}(\alpha+M).$$

因此，(1)式所给出的定义在逻辑上是没有矛盾的.

我们现在证明 $\bar{\boldsymbol{A}}$ 是 V/M 内的一个线性变换：

(i) $$\begin{aligned}\bar{\boldsymbol{A}}[(\alpha+M)+(\beta+M)]&=\bar{\boldsymbol{A}}[(\alpha+\beta)+M]\\&=\boldsymbol{A}(\alpha+\beta)+M=(\boldsymbol{A}\alpha+\boldsymbol{A}\beta)+M\\&=(\boldsymbol{A}\alpha+M)+(\boldsymbol{A}\beta+M)\\&=\bar{\boldsymbol{A}}(\alpha+M)+\bar{\boldsymbol{A}}(\beta+M);\end{aligned}$$

(ii) $$\begin{aligned}\bar{\boldsymbol{A}}[k(\alpha+M)]&=\bar{\boldsymbol{A}}[k\alpha+M]=\boldsymbol{A}(k\alpha)+M\\&=k\boldsymbol{A}\alpha+M=k(\boldsymbol{A}\alpha+M)=k\bar{\boldsymbol{A}}(\alpha+M).\end{aligned}$$

定义 由(1)式所定义的 V/M 内的线性变换 $\bar{\boldsymbol{A}}$ 称为 $\boldsymbol{A}$ 在商空间 V/M 内的**诱导变换**.

现在在 M 内取一组基 $\varepsilon_1,\cdots,\varepsilon_r$，扩充成 V 的一组基

$$\varepsilon_1,\cdots,\varepsilon_r,\varepsilon_{r+1},\cdots,\varepsilon_n.$$

设

$$\boldsymbol{A}\varepsilon_{r+k}=a_{r+1\,r+k}\varepsilon_{r+1}+\cdots+a_{n\,r+k}\varepsilon_n+\beta_{r+k}$$
$$(k=1,2,\cdots,n-r),$$

其中 $\beta_{r+k}\in M$. 在第四章 § 4 中已指出

$$\varepsilon_{r+1}+M,\ \varepsilon_{r+2}+M,\ \cdots,\ \varepsilon_n+M$$

组成 V/M 的一组基. 为简便计，令 $\varepsilon_{r+k}+M=\bar{\varepsilon}_{r+k}$，有

$$\begin{aligned}\bar{\boldsymbol{A}}\bar{\varepsilon}_{r+k}&=\boldsymbol{A}(\varepsilon_{r+k}+M)=(a_{r+1\,r+k}\varepsilon_{r+1}+\cdots+a_{n\,r+k}\varepsilon_n)+M\\&=a_{r+1\,r+k}\bar{\varepsilon}_{r+1}+\cdots+a_{n\,r+k}\bar{\varepsilon}_n.\end{aligned}$$

这表明 $\bar{\boldsymbol{A}}$ 在基 $\bar{\varepsilon}_{r+1},\bar{\varepsilon}_{r+2},\cdots,\bar{\varepsilon}_n$ 下的矩阵为

$$\bar{A}=\begin{pmatrix}a_{r+1\,r+1}&\cdots&a_{r+1\,n}\\\vdots&&\vdots\\a_{n\,r+1}&\cdots&a_{n\,n}\end{pmatrix}.$$

注意 $\boldsymbol{A}$ 在 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 这组基下的矩阵具有如下分块形式

$$A=\begin{pmatrix} A_1 & A_2 \\ 0 & \bar{A} \end{pmatrix}.$$

$\boldsymbol{A}$ 的特征多项式为

$$f(\lambda)=|\lambda E-A|=|\lambda E-A_1|\cdot|\lambda E-\bar{A}|,$$

故 $\bar{\boldsymbol{A}}$ 的特征多项式 $|\lambda E-\bar{A}|$ 为 $\boldsymbol{A}$ 的特征多项式的因子. 从而, $\bar{\boldsymbol{A}}$ 的特征值为 $\boldsymbol{A}$ 的特征值的一部分.

习 题 五

1. 设 $\boldsymbol{A}$ 是线性空间 V 内的一个线性变换, M,N 是 $\boldsymbol{A}$ 的两个不变子空间. 证明: $M+N$ 与 $M\cap N$ 都是 $\boldsymbol{A}$ 的不变子空间.

2. 设 $\boldsymbol{A}$ 是 n 维线性空间 V 内的一个线性变换, 在 V 的一组基下其矩阵成若当块

$$J=\begin{pmatrix} \lambda_0 & 1 & & \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0 \end{pmatrix},$$

证明: 当 $n>1$ 时, 对 $\boldsymbol{A}$ 的任一非平凡不变子空间 M, 都不存在 $\boldsymbol{A}$ 的不变子空间 N, 使

$$V=M\oplus N.$$

3. 设 V 是实数域上的二维线性空间. 线性变换 $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2$ 下的矩阵为

$$A=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \quad (\theta\neq k\pi),$$

证明 $\boldsymbol{A}$ 没有非平凡不变子空间.

*4. 设 V 是数域 K 上的二维线性空间, F 为 V 内在基 $\varepsilon_1,\varepsilon_2$ 下有矩阵

$$\begin{pmatrix} 0 & a \\ 1-a & 0 \end{pmatrix} \quad (a\in K)$$

的线性变换所成的集合, 证明 F 中的线性变换没有公共非平凡不

变子空间.

*5. 设 V 是实数域上的一个 n 维线性空间,$\boldsymbol{A}$ 是 V 内的一个线性变换. 证明 $\boldsymbol{A}$ 必有一个一维或二维的不变子空间.

6. 设 $\boldsymbol{A},\boldsymbol{B}$ 是 n 维线性空间 V 内两个线性变换,且 $\boldsymbol{AB}=\boldsymbol{BA}$. λ 是 $\boldsymbol{A}$ 的一个特征值,V_λ 是属于特征值 λ 的特征子空间. 证明 V_λ 是 $\boldsymbol{B}$ 的不变子空间.

*7. 设 $\boldsymbol{A}$ 是 n 维线性空间 V 内的一个线性变换,且在 V 内存在一组基,使 $\boldsymbol{A}$ 在此组基下的矩阵成对角形. 设 M 是 $\boldsymbol{A}$ 的一个不变子空间,证明 M 内存在一组基,使 $\boldsymbol{A}|_M$ 在该组基下的矩阵成对角形.

*8. 设 V 是数域 K 上的 n 维线性空间,$\boldsymbol{A},\boldsymbol{B}$ 是 V 内两个线性变换,且 $\boldsymbol{AB}=\boldsymbol{BA}$. 如果 $\boldsymbol{A},\boldsymbol{B}$ 的矩阵都可对角化,证明 V 内存在一组基,使 $\boldsymbol{A},\boldsymbol{B}$ 在该组基下的矩阵同时成对角形.

第六章　双线性函数与二次型

在第五章中所研究的线性变换，是数域上一个线性空间 V 到自身的一个满足线性条件的映射. 在这一章中，我们将要来研究 V 到数域 K 的一个映射 f. 它可以看作以 V 为定义域，而以 K 的一个子集作为值域的一个函数. 我们将首先研究最简单的满足线性条件的函数，在这个基础上，再进而讨论稍复杂一点的函数，即二次齐次函数——二次型.

§1　双线性函数

线性函数的概念

定义　设 V 是数域 K 上的一个线性空间. 如果对 V 中任一向量 α，都按某一法则 f 对应于数域 K 内唯一确定的一个数，记作 $f(\alpha)$，而且有：

(i) 对任意 $\alpha,\beta\in V$，有 $f(\alpha+\beta)=f(\alpha)+f(\beta)$；

(ii) 对任意 $k\in K,\alpha\in V$，有 $f(k\alpha)=kf(\alpha)$，

则称 $f(\alpha)$ 为 V 上的一个**线性函数**.

定义中的条件(i)，(ii)与下面的条件等价：

(iii) 对任意 $k,l\in K,\alpha,\beta\in V$，有

$$f(k\alpha+l\beta)=kf(\alpha)+lf(\beta).$$

这个定义与数学分析中函数的定义类似. 只是现在“自变量”是 V 中的向量，不是实数. 但我们可以把它理解为以向量作“自变量”的函数.

这个定义和上一章中线性变换的定义也有某些类似之处. 它

们的不同点是，线性变换 $\boldsymbol{A}$ 作用在 $\alpha\in V$ 上后，$\boldsymbol{A}\alpha$ 仍是 V 中的向量，而线性函数 f 作用在 α 上后，$f(\alpha)$ 是数域 K 中的数，不是 V 中的向量.

现设 V 是 n 维线性空间，在 V 内取定一组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n,$$

任一向量 $\alpha\in V$ 可表为

$$\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n \quad (x_i\in K,i=1,2,\cdots,n).$$

对于 V 上任一线性函数 $f(\alpha)$，显然有

$$f(\alpha)=x_1f(\varepsilon_1)+x_2f(\varepsilon_2)+\cdots+x_nf(\varepsilon_n).$$

只要知道了 $f(\varepsilon_1),f(\varepsilon_2),\cdots,f(\varepsilon_n)$，那么 $f(\alpha)$ 就被完全确定了. 因此，对于 V 上的线性函数，我们有类似于线性变换的两条结论：

(1) V 上任一线性函数 $f(\alpha)$ 由它在一组基下的函数值 $f(\varepsilon_1),f(\varepsilon_2),\cdots,f(\varepsilon_n)$ 唯一确定. 换句话说，如果 $f(\alpha),g(\alpha)$ 是 V 上两个线性函数，且 $f(\varepsilon_i)=g(\varepsilon_i)(i=1,2,\cdots,n)$，则 $f(\alpha)\equiv g(\alpha)$；

(2) 任给 K 内 n 个数 $a_1,a_2,\cdots,a_n$，在 V 上存在唯一的一个线性函数 $f(\alpha)$，使

$$f(\varepsilon_i)=a_i \quad (i=1,2,\cdots,n).$$

这两条结论的证明方法与线性变换的相应命题证法一样，留给读者作为练习.

对一个线性函数 $f(\alpha)$，如果设

$$f(\varepsilon_1)=a_1,\ f(\varepsilon_2)=a_2,\ \cdots,\ f(\varepsilon_n)=a_n;$$

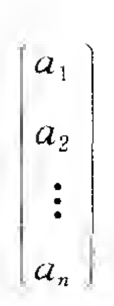
$$\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n.$$

则线性函数 $f(\alpha)$ 可以表成如下形式：

$$f(\alpha)=x_1a_1+x_2a_2+\cdots+x_na_n$$

$$=(x_1\ \ x_2\ \ \cdots\ \ x_n)\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}.$$

列矩阵



$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

称为线性函数 $f(\alpha)$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵. 显然,取定 V 内一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 后,我们可以在 V 上全体线性函数所组成的集合和 K 上全体 $n\times 1$ 矩阵所成的集合之间建立起一一对应的关系. 这与线性变换和它的矩阵之间的关系类似.

*对偶空间

设 V 是数域 K 上的一个线性空间,V^* 是 V 上全体线性函数所组成的集合. 在 V^* 内定义加法与数乘运算如下:

(1) 加法. 对任意 $f,g\in V^*$,定义 V 到 K 的一个映射(即 V 上的一个函数)如下:

$$(f+g)(\alpha)=f(\alpha)+g(\alpha).$$

容易验证 $f+g$ 也是 V 上的一个线性函数,称为 f 与 g 的**和**;

(2) 数乘. 对任意 $k\in K,f\in V^*$,定义 V 到 K 的一个映射如下:

$$(kf)(\alpha)=kf(\alpha).$$

容易验证,kf 也是 V 上的一个线性函数,称它为 k 与 f 的**数乘**.

通过直接验证可知上面定义的加法与数乘运算满足线性空间的八条公理,所以 V^* 关于上述加法与数乘运算组成数域 K 上的一个线性空间,这个空间称为 V 的**对偶空间**或**共轭空间**.

现在设 V 是 n 维线性空间,在 V 内取定一组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n.$$

我们定义 V 上 n 个线性函数如下:

$$f_i(\varepsilon_j)=\delta_{ij}\quad(i=1,2,\cdots,n;j=1,2,\cdots,n)\tag{1}$$

(只要给定 f_i 在 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 处的值,f_i 就被唯一确定了). 现在我们证明这 n 个线性函数

$$f_1(\alpha),f_2(\alpha),\cdots,f_n(\alpha)$$

组成 V^* 的一组基.

1. $f_1,f_2,\cdots,f_n$ 线性无关. 设

$$k_1f_1(\alpha)+k_2f_2(\alpha)+\cdots+k_nf_n(\alpha)\equiv 0,$$

上式表示以 V 中任一 α 代入等式均成立. 现以 $\alpha=\varepsilon_j$ 代入,由(1)式可知有

$$k_j=k_jf_j(\varepsilon_j)=0 \quad (j=1,2,\cdots,n).$$

这就证明了 $f_1,f_2,\cdots,f_n$ 是线性无关的.

2. 再证任一 $f\in V^*$ 均可被 $f_1,f_2,\cdots,f_n$ 线性表示. 设

$$f(\varepsilon_1)=a_1,\ f(\varepsilon_2)=a_2,\ \cdots,\ f(\varepsilon_n)=a_n.$$

考察

$$\tilde{f}(\alpha)=a_1f_1(\alpha)+a_2f_2(\alpha)+\cdots+a_nf_n(\alpha).$$

因为 V^* 是一线性空间,V^* 内的向量 $f_1,f_2,\cdots,f_n$ 的线性组合 $\tilde{f}\in V^*$. 再由(1)式,有

$$\tilde{f}(\varepsilon_j)=a_jf_j(\varepsilon_j)=a_j=f(\varepsilon_j) \quad (j=1,2,\cdots,n).$$

$\tilde{f}(\alpha)\equiv f(\alpha)$,故

$$f(\alpha)=a_1f_1(\alpha)+a_2f_2(\alpha)+\cdots+a_nf_n(\alpha).$$

定义 在 V 内取定一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,则由(1)式所定义的 $f_1,f_2,\cdots,f_n$ 构成 V^* 的一组基,称这组基为 V 内 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 这组基的**对偶基**.

从上面的分析可知,V^* 也是数域 K 上的 n 维线性空间. 因为 $\dim V^*=\dim V$,所以 V^* 和 V 是同构的线性空间.

双线性函数

定义 设 V 是数域 K 上的线性空间. 如果 V 中任意一对有序向量(α,β)都按照某一法则 f 对应于 K 内唯一确定的一个数,记作 $f(\alpha,\beta)$,且

(i) 对任意 $k_1,k_2\in K,\alpha_1,\alpha_2,\beta\in V$,有

$$f(k_1\alpha_1+k_2\alpha_2,\beta)=k_1f(\alpha_1,\beta)+k_2f(\alpha_2,\beta);$$

(ⅱ) 对任意 $l_1,l_2\in K,\alpha,\beta_1,\beta_2\in V$，有

$$f(\alpha,l_1\beta_1+l_2\beta_2)=l_1f(\alpha,\beta_1)+l_2f(\alpha,\beta_2).$$

则称 $f(\alpha,\beta)$是 V 上的一个**双线性函数**.

根据这个定义可知：如令 β 保持不动，则 $f(\alpha,\beta)$是 α 的线性函数；同样，如令 α 保持不动，则 $f(\alpha,\beta)$是 β 的线性函数. 这就是称它为双线性函数的原因. 双线性函数可以看作是以 V 内向量为自变量的二元线性函数.

现设 V 是 n 维线性空间. 在 V 内取定一组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n.$$

又设

$$\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n;$$

$$\beta=y_1\varepsilon_1+y_2\varepsilon_2+\cdots+y_n\varepsilon_n.$$

那么，按定义，有

$$\begin{aligned}f(\alpha,\beta)&=f\left(\sum_{i=1}^{n}x_i\varepsilon_i,\sum_{j=1}^{n}y_j\varepsilon_j\right)\\&=\sum_{i=1}^{n}x_if\left(\varepsilon_i,\sum_{j=1}^{n}y_j\varepsilon_j,\right)\\&=\sum_{i=1}^{n}\sum_{j=1}^{n}x_iy_jf(\varepsilon_i,\varepsilon_j).\end{aligned}$$

由此可见，$f(\alpha,\beta)$由它在一组基处的函数值唯一确定. 显然，现在有类似于线性变换的两条结论：

(1) V 上一个双线性函数 $f(\alpha,\beta)$由它在一组基处的函数值 $f(\varepsilon_i,\varepsilon_j)(i,j=1,2,\cdots,n)$唯一确定. 换句话说，如果有一个双线性函数 $g(\alpha,\beta)$满足 $g(\varepsilon_i,\varepsilon_j)=f(\varepsilon_i,\varepsilon_j)$，则 $g(\alpha,\beta)\equiv f(\alpha,\beta)$；

(2) 任给一个 n 阶方阵

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}\quad(a_{ij}\in K),$$

必存在 V 上唯一的一个双线性函数 $f(\alpha,\beta)$，使

$$f(\varepsilon_i,\varepsilon_j)=a_{ij}\quad(i,j=1,2,\cdots,n).$$

这两条结论的证明与线性变换的对应结论的证明类似，留给读者作为练习.

我们称

$$A=\begin{pmatrix} f(\varepsilon_1,\varepsilon_1) & f(\varepsilon_1,\varepsilon_2) & \cdots & f(\varepsilon_1,\varepsilon_n) \\ f(\varepsilon_2,\varepsilon_1) & f(\varepsilon_2,\varepsilon_2) & \cdots & f(\varepsilon_2,\varepsilon_n) \\ \vdots & \vdots & & \vdots \\ f(\varepsilon_n,\varepsilon_1) & f(\varepsilon_n,\varepsilon_2) & \cdots & f(\varepsilon_n,\varepsilon_n) \end{pmatrix}$$

为双线性函数 $f(\alpha,\beta)$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵. 这样，V 上每个双线性函数对应于数域 K 上的一个 n 阶方阵. 根据(1)与(2)，这个对应是 V 上全体双线性函数所成的集合和 $M_n(K)$ 之间的一个一一对应.

设 $f(\varepsilon_i,\varepsilon_j)=a_{ij}$，于是 $f(\alpha,\beta)$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵 $A=(a_{ij})$. 利用矩阵乘法，双线性函数 $f(\alpha,\beta)$ 可表作

$$\begin{aligned} f(\alpha,\beta) &= \sum_{i=1}^{n}\sum_{j=1}^{n} f(\varepsilon_i,\varepsilon_j)x_iy_j = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_iy_j \\ &= (x_1\ \ x_2\ \ \cdots\ \ x_n)A\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}. \end{aligned}$$

如令

$$X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},\quad Y=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

则

$$f(\alpha,\beta)=X'AY.$$

反过来，如果一个双线性函数被表示成上述矩阵乘积的形式，那

么，只要计算一下 $f(\varepsilon_i,\varepsilon_j)$就可以知道 A 就是 $f(\alpha,\beta)$在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵.

双线性函数在不同基下的矩阵

我们已经知道，一个线性变换在不同的两组基下的矩阵是相似的. 我们现在来指出：一个双线性函数在不同的两组基下的矩阵也有与此相类似的关系.

设 $f(\alpha,\beta)$为 V 内一个双线性函数. 在 V 内取定两组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n;$$

$$\eta_1,\eta_2,\cdots,\eta_n.$$

设

$$f(\varepsilon_i,\varepsilon_j)=a_{ij},\quad f(\eta_i,\eta_j)=b_{ij}\quad (i,j=1,2,\cdots,n).$$

记 $A=(a_{ij}),B=(b_{ij})$，则 A,B 分别是 $f(\alpha,\beta)$在这两组基下的矩阵. 令

$$\begin{cases}\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X,\\ \beta=y_1\varepsilon_1+y_2\varepsilon_2+\cdots+y_n\varepsilon_n=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)Y.\end{cases}$$

$$\begin{cases}\alpha=\overline{x}_1\eta_1+\overline{x}_2\eta_2+\cdots+\overline{x}_n\eta_n=(\eta_1,\eta_2,\cdots,\eta_n)\overline{X},\\ \beta=\overline{y}_1\eta_1+\overline{y}_2\eta_2+\cdots+\overline{y}_n\eta_n=(\eta_1,\eta_2,\cdots,\eta_n)\overline{Y}.\end{cases}$$

又设两组基之间的过渡矩阵为 T

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T.$$

于是有坐标变换公式

$$X=T\overline{X};\quad Y=T\overline{Y}.$$

将上面的式子代入 $f(\alpha,\beta)$的矩阵表达式中，有

$$\begin{aligned}f(\alpha,\beta)&=X'AY=(T\overline{X})'A(T\overline{Y})\\&=\overline{X}'(T'AT)\overline{Y}.\end{aligned}$$

上式表示 $f(\alpha,\beta)$在基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的矩阵为 $T'AT$(参看上一段最后的说明)，但已知它在这组基下有矩阵 B，且双线性函数在一组基下的矩阵是唯一的，故有

$$B=T'AT.$$

这就是同一个双线性函数 $f(\alpha,\beta)$ 在两组不同的基下的矩阵之间的关系.

定义 给定数域 K 上两个 n 阶方阵 A,B. 如果存在 K 上一个可逆的 n 阶方阵 T,使 $B=T'AT$,则称 B 与 A **合同**.

因此,同一个双线性函数在不同基下的矩阵是合同的. 反过来,两个互相合同的矩阵可以看作是某一个双线性函数在两组不同基下的矩阵,其证明方法与线性变换的相应命题(第五章命题2.3)相同,请读者自己证明.

方阵间的合同关系是矩阵之间的又一重要关系. 它与矩阵的相似关系不同,但形式上类似. 显然,矩阵之间的合同关系也具有如下几条基本性质

(1) 反身性: $A=E'AE$;

(2) 对称性: 若 $B=T'AT$,则 $A=(T^{-1})'BT^{-1}$;

(3) 传递性: 若 $A=T_1'BT_1, B=T_2'CT_2$,则

$$A=T_1'BT_1=T_1'(T_2'CT_2)T_1=(T_2T_1)'C(T_2T_1).$$

因此,矩阵的合同是一个等价关系. 在此等价关系下的等价类称作合同类. 我们自然也会想从每个合同类中挑选出一个最简单的矩阵(最好是对角矩阵)来作为该合同类的代表. 使用双线性函数的语言,就是对每一个双线性函数 $f(\alpha,\beta)$,要设法在 V 内找出一组基,使 $f(\alpha,\beta)$ 在这组基下的矩阵具有最简单的形式.

在本书中,我们仅限于对一类最常用的双线性函数来讨论这个问题. 这就是下面所要讲的对称双线性函数.

对称双线性函数

定义 设 V 是数域 K 上的线性空间,$f(\alpha,\beta)$ 是 V 上的一个双线性函数. 如果对任意 $\alpha,\beta\in V$,有 $f(\alpha,\beta)=f(\beta,\alpha)$,则称 $f(\alpha,\beta)$ 是一个**对称双线性函数**.

显然,有限维线性空间 V 内的对称双线性函数在一组基下的矩阵一定具有它们自己的特点. 在阐述这一特点之前,首先给出如

下的概念.

定义 设 $A=(a_{ij})$是数域 K 上的一个 n 阶方阵. 如果 $A'=A$,即

$$a_{ij}=a_{ji} \quad (i,j=1,2,\cdots,n),$$

则称 A 是一对**对称方阵**. 如果 $A'=-A$,即

$$a_{ij}=-a_{ji} \quad (i,j=1,2,\cdots,n)$$

(由此即知 $a_{ii}=0,i=1,2,\cdots,n$),则称 A 是一个**反对称方阵**.

不难看出,对称双线性函数在任一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵 $(f(\varepsilon_i,\varepsilon_j))$都是对称矩阵(因为 $f(\varepsilon_i,\varepsilon_j)=f(\varepsilon_j,\varepsilon_i)$). 反之,任给一个对称矩阵 $A=(a_{ij})$(其中 $a_{ij}=a_{ji}$),定义 V 内一个双线性函数 $f(\alpha,\beta)$,使满足

$$f(\varepsilon_i,\varepsilon_j)=a_{ij} \quad (i,j=1,2,\cdots,n).$$

因为 $f(\varepsilon_i,\varepsilon_j)=a_{ij}=a_{ji}=f(\varepsilon_j,\varepsilon_i)$,故对任意,$\alpha,\beta\in V$,有

$$f(\alpha,\beta)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_iy_j=\sum_{j=1}^{n}\sum_{i=1}^{n}a_{ji}y_jx_i=f(\beta,\alpha).$$

即 $f(\alpha,\beta)$是 V 上一个对称双线性函数,它在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵即为 A. 因此,给定了 V 内一组基之后,就在 V 上全体对称双线性函数所成的集合和 K 上全体 n 阶对称方阵所成的集合之间建立了一一对应的关系.

下面是关于对称双线性函数的基本定理.

定理 1 设 V 是数域 K 上的 n 维线性空间,$f(\alpha,\beta)$是 V 上的一个对称双线性函数,则在 V 内存在一组基,使 $f(\alpha,\beta)$在这组基下的矩阵成对角形.

证 对 V 的维数 n 作数学归纳法.

$n=1$ 时定理是显然的. 设对 $n-1$ 维线性空间定理成立,证明对 n 维线性空间定理也成立.

首先,若 $f(\alpha,\beta)\equiv0$,定理自然成立. 如果不是这种情况,则不能对一切 $\alpha\in V$ 都有 $f(\alpha,\alpha)=0$. 因为如果这样,则对任意的 $\alpha,\beta\in V$,有

$$0=f(\alpha+\beta,\alpha+\beta)=f(\alpha,\alpha)+2f(\alpha,\beta)+f(\beta,\beta)$$
$$=2f(\alpha,\beta),$$

即对一切 α,β 有 $f(\alpha,\beta)=0$，与假设矛盾.

于是，可在 V 内取定 ε_1，使 $f(\varepsilon_1,\varepsilon_1)=d_1\neq 0$. 把 ε_1 扩充成 V 的一组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n.$$

令
$$\begin{cases}\eta_1=\varepsilon_1,\\ \eta_i'=\dfrac{f(\varepsilon_1,\varepsilon_i)}{d_1}\varepsilon_1+\varepsilon_i \quad (i=2,3,\cdots,n).\end{cases}$$

显然，$\eta_1,\eta_2',\cdots,\eta_n'$ 是 V 的一组基，且

$$f(\eta_1,\eta_i')=0 \quad (i=2,3,\cdots,n).$$

命 $M=L(\eta_2',\cdots,\eta_n')$. 这是一个 $n-1$ 维线性空间，$f(\alpha,\beta)$ 可以看作 M 内的对称双线性函数. 按归纳假设，在 M 内存在一组 $\eta_2,\cdots,\eta_n$，使 $f(\alpha,\beta)$ 在这组基下的矩阵成对角形. 即有

$$f(\eta_i,\eta_j)=d_i\delta_{ij} \quad (i,j=2,3,\cdots,n).$$

因为 $\eta_1,\eta_2,\cdots,\eta_n$ 与 $\eta_1',\eta_2',\cdots,\eta_n'$ 等价，故它也是 V 的一组基，且

$$f(\eta_i,\eta_j)=d_i\delta_{ij} \quad (i,j=1,2,\cdots,n).$$

于是 $f(\alpha,\beta)$ 在这组基下的矩阵成对角形. ▌

推论 设 A 是数域 K 上的一个 n 阶对称方阵，则存在 K 上的一个可逆矩阵 T，使 $T'AT=D$ 为对角形.

证 把 A 看作 K 上 n 维线性空间 V 内对称双线性函数 $f(\alpha,\beta)$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵. 由定理 1，在 V 内存在一组基 $\eta_1,\eta_2,\cdots,\eta_n$，使 $f(\alpha,\beta)$ 在这组基下的矩阵成对角形 D. 设

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T,$$

则 $T'AT=D$. ▌

最后，我们介绍双线性函数的秩的概念. 我们已经知道，同一个双线性函数在不同基下的矩阵是互相合同的，而互相合同的矩阵秩相同（这是因为：根据第二章命题 2,3，一个矩阵左乘或右乘一个满秩方阵后，其秩不变）. 因此，可以有如下的定义：

定义 设$f(\alpha,\beta)$是n维线性空间V内的一个双线性函数，它在某一组基下的矩阵A的秩$r(A)$称为$f(\alpha,\beta)$的**秩**. 如果A是满秩的，即$r(A)=n$，则称$f(\alpha,\beta)$是**满秩双线性函数**(或称**非退化双线性函数**).

习　题　一

1. 设V是数域K上的n维线性空间，$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是$V$的一组基，$a_1,a_2,\cdots,a_n$为$K$内的$n$个数. 证明：在$V$上存在唯一的一个线性函数$f(\alpha)$，满足

$$f(\varepsilon_i)=a_i \quad (i=1,2,\cdots,n).$$

2. 设$f(\alpha)$是线性空间V内的一个线性函数. 证明：

$$f(0)=0,\ f(-\alpha)=-f(\alpha).$$

3. 在实数域上线性空间$C[a,b]$上定义函数$I(f)$如下：对$f(x)\in C[a,b]$，令

$$I(f)=\int_a^b f(x)\mathrm{d}x.$$

证明：$I(f)$是$C[a,b]$上的一个线性函数.

4. 在线性空间$M_n(K)$上定义函数如下：

$$f(A)=\mathrm{Tr}(A).$$

证明：$f(A)$是$M_n(K)$上的一个线性函数.

5. 在三维几何空间中全体向量组成的实数域上线性空间上定义二元函数如下：

$$f(\vec{a},\vec{b})=\vec{a}\cdot\vec{b}.$$

证明这是一个双线性函数.

6. 在实数域上线性空间$C[a,b]$上定义二元函数如下：对$f(x),g(x)\in C[a,b]$，令

$$I(f,g)=\int_a^b f(x)g(x)\mathrm{d}x.$$

证明这是一个双线性函数.

7. 在 K^n 中定义函数如下：若

$$\alpha=(a_1,a_2,\cdots,a_n),\quad \beta=(b_1,b_2,\cdots,b_n),$$

则令

$$f(\alpha,\beta)=a_1b_1+a_2b_2+\cdots+a_nb_n.$$

证明这是一个双线性函数.

8. 在 K^4 中定义函数如下：若

$$\alpha=(x_1,x_2,x_3,x_4),\quad \beta=(y_1,y_2,y_3,y_4),$$

则令

$$f(\alpha,\beta)=3x_1y_2-5x_2y_1+x_3y_4-4x_4y_3.$$

(1) 证明 $f(\alpha,\beta)$是一个双线性函数.

(2) 在 K^4 中给定一组基

$$\varepsilon_1=(1,2,-1,0),\quad \varepsilon_2=(1,-1,1,1),$$
$$\varepsilon_3=(-1,2,1,1),\quad \varepsilon_4=(-1,-1,0,1),$$

求 $f(\alpha,\beta)$在这组基下的矩阵.

(3) 在 K^4 内另给一组基 $\eta_1,\eta_2,\eta_3,\eta_4$，且

$$(\eta_1,\eta_2,\eta_3,\eta_4)=(\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4)T,$$

其中

$$T=\begin{pmatrix}1&1&1&1\\1&1&-1&-1\\1&-1&1&-1\\1&-1&-1&1\end{pmatrix},$$

求 $f(\alpha,\beta)$在基 $\eta_1,\eta_2,\eta_3,\eta_4$ 下的矩阵.

9. 在 $M_n(K)$内定义函数如下：

$$f(A,B)=\mathrm{Tr}(AB).$$

(1) 证明 $f(A,B)$是一个对称双线性函数.

(2) 令 $n=2$，在 $M_2(K)$内取一组基

$$\varepsilon_{11}=\begin{pmatrix}1&0\\0&0\end{pmatrix},\ \varepsilon_{12}=\begin{pmatrix}0&1\\0&0\end{pmatrix},$$

$$\varepsilon_{21}=\begin{pmatrix}0&0\\1&0\end{pmatrix},\ \varepsilon_{22}=\begin{pmatrix}0&0\\0&1\end{pmatrix},$$

求 $f(A,B)$在这组基下的矩阵.

(3) 在 $M_2(K)$内另取一组基

$$\eta_1=\begin{pmatrix}1&0\\0&1\end{pmatrix},\ \eta_2=\begin{pmatrix}1&0\\0&-1\end{pmatrix},$$

$$\eta_3=\begin{pmatrix}0&1\\1&0\end{pmatrix},\ \eta_4=\begin{pmatrix}0&1\\-1&0\end{pmatrix},$$

求出两组基之间的过渡矩阵 T

$$(\eta_1,\eta_2,\eta_3,\eta_4)=(\varepsilon_{11},\varepsilon_{12},\varepsilon_{21},\varepsilon_{22})T.$$

再求 $f(A,B)$在基 $\eta_1,\eta_2,\eta_3,\eta_4$ 下的矩阵.

(4) 在 $n=2$ 在情况下求 $f(A,B)$的秩.

10. 设 V 是数域 K 上的 n 维线性空间,$f(\alpha,\beta)$是 V 上的一个双线性函数. 证明:$f(\alpha,\beta)$满秩的充分必要条件是:当对一切 $\beta\in V$ 有 $f(\alpha,\beta)=0$ 时,必定有 $\alpha=0$.

11. 证明第 9 题中的双线性函数 $f(A,B)$是满秩的.

12. 在 $\mathscr{R}^4$ 中定义函数如下:若

$$\alpha=(x_1,x_2,x_3,x_4),\quad \beta=(y_1,y_2,y_3,y_4),$$

则令

$$f(\alpha,\beta)=x_1y_1+x_2y_2+x_3y_3-x_4y_4.$$

(1) 证明 $f(\alpha,\beta)$是对称双线性函数;

(2) 求 $f(\alpha,\beta)$在基

$$\varepsilon_1=(1,0,0,0),\quad \varepsilon_2=(0,1,0,0),$$

$$\varepsilon_3=(0,0,1,0),\quad \varepsilon_4=(0,0,0,1)$$

下的矩阵;

(3) 求 $f(\alpha,\beta)$在基

$$\eta_1=(2,1,-1,1),\quad \eta_2=(0,3,1,0),$$

$$\eta_3=(5,3,2,1),\quad \eta_4=(6,6,1,3)$$

下的矩阵;

（4） 证明 $f(\alpha,\beta)$是满秩的；

（5） 求一个向量 $\alpha\neq 0$，使 $f(\alpha,\alpha)=0$.

13. 在 K^4 内给定双线性函数 $f(\alpha,\beta)$如下，试判断哪些是对称的，哪些是满秩的. 设

$$\alpha=(x_1,x_2,x_3,x_4),\ \beta=(y_1,y_2,y_3,y_4).$$

（1） $f(\alpha,\beta)=x_1y_2-x_1y_3+2x_1y_4-x_2y_1+x_2y_4-x_3y_1$
$-2x_3y_4-2x_4y_1-x_4y_2+2x_4y_3$；

（2） $f(\alpha,\beta)=x_1y_1+2x_1y_2+2x_2y_1+4x_2y_2-x_3y_3$
$-x_3y_4-x_4y_3+x_4y_4$；

（3） $f(\alpha,\beta)=-x_1y_3+x_1y_4+x_3y_1+x_3y_4-x_4y_1-x_4y_3$；

（4） $f(\alpha,\beta)=x_1y_4+x_4y_1$.

14. 证明：

$$\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix} 与 \begin{pmatrix}\lambda_{i_1} & & & \\ & \lambda_{i_2} & & \\ & & \ddots & \\ & & & \lambda_{i_n}\end{pmatrix}$$

合同，其中 $i_1,i_2,\cdots,i_n$ 是 $1,2,\cdots,n$ 的一个排列.

15. 设 A 是一个 n 阶方阵，证明：

（1） A 反对称当且仅当对任一 n 维列向量 X，有

$$X'AX=0;$$

（2） 若 A 对称，且对任一 n 维列向量 X 有 $X'AX=0$，那么 $A=0$.

§2 二次型和它的标准形

在这一节里，我们将把§1 中所得的结果应用到 n 个变量的二次齐次函数上来，所得的结果在理论上和实际工作中都有着重要的应用. 我们首先介绍与此有关的一些基本概念.

二次型的基本概念

设 V 是数域 K 上的 n 维线性空间，$f(\alpha,\beta)$是 V 内一个对称双线性函数. 定义 V 内一个新函数 $Q_f(\alpha)=f(\alpha,\alpha)$，称为 V 内一个**二次型函数**.

因为 $f(\alpha,\beta)=f(\beta,\alpha)$，所以

$$f(\alpha+\beta,\alpha+\beta)=f(\alpha,\alpha)+2f(\alpha,\beta)+f(\beta,\beta),$$

由此得二次型函数与对称双线性函数的如下基本关系：

$$f(\alpha,\beta)=\frac{1}{2}\{Q_f(\alpha+\beta)-Q_f(\alpha)-Q_f(\beta)\}.$$

如果 V 内另有对称双线性函数 $g(\alpha,\beta)$，其对应的二次型函数 $Q_g(\alpha)\equiv Q_f(\alpha)$，则由上述基本关系式立即推出 $f(\alpha,\beta)\equiv g(\alpha,\beta)$. 所以一个对称双线性函数由它对应的二次型函数唯一决定.

在 V 内取定一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，又设

$$\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X,$$

那么，按上一节的结果，我们有

$$Q_f(\alpha)=f(\alpha,\alpha)=X'AX=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j,$$

这里 $A=(a_{ij})=(f(\varepsilon_i,\varepsilon_j))$是对称双线性函数 $f(\alpha,\beta)$在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵，它是数域 K 上一个 n 阶对称矩阵. 如果把 $x_1,x_2,\cdots,x_n$ 看作 n 个自变量，那么 $Q_f(\alpha)$的上述表达式就是数学分析中研究的 n 元函数的一个特例. 当然，数学分析中自变量的定义域是实数域 $\mathscr{R}$，而这里自变量 $x_i(i=1,2,\cdots,n)$的定义域是一般的数域 K，这一点是不同的.

定义 以数域 K 的元素作系数的 n 个变量 $x_1,x_2,\cdots,x_n$ 的二次齐次函数

$$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j \quad (a_{ij}=a_{ji}) \tag{1}$$

称为数域 K 上的一个**二次型**，其系数所成的矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

称为此**二次型的矩阵**. A 的秩 $r(A)$ 称为此**二次型的秩**.

令

$$X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

二次型(1)可以表成矩阵乘积的形状

$$f=X'AX,$$

这已经在上面 $Q_f(\alpha)$ 的表达式中指明了.

所以,数域 K 上一个二次型 f 就是 V 内一个二次型函数 $Q_f(\alpha)$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的解析表达式. 这样,上一节关于对称双线性函数所得的结果可以直接用到二次型 f 上来.

二次型在数学各领域以至自然科学、工程技术中都有广泛的应用. 例如,在理论力学中讨论质点系的动能、位能时,就需要二次型的概念. 本节的任务,是要从代数的观点来阐述二次型的基本理论.

在实际工作中碰到的二次型,其表达式中 x_ix_j 与 x_jx_i 两项可能是合并在一起的,为了用上一节所得到的结果,我们必须把 f 与 V 内对应的二次型函数 $Q_f(\alpha)$ 或对应的对称双线性函数 $f(\alpha,\beta)$ 联系起来,这样就应当把该项系数平分为二,得出两项 $a_{ij}x_ix_j$ 及 $a_{ji}x_jx_i$,且 $a_{ji}=a_{ij}$. 然后再写出此二次型的矩阵 $A=(a_{ij})$.

例 1 给定二次型

$$f=x_1^2-2x_1x_2+3x_1x_3-x_2x_3+4x_3^2.$$

为了写出它的矩阵,我们把它写成

$$
\begin{aligned}
f= \quad & x_1^2 \quad -x_1x_2 \quad + \frac{3}{2}x_1x_3 \\
& - \ x_2x_1+0\cdot x_2^2 \ -\frac{1}{2}x_2x_3 \\
& +\frac{3}{2}x_3x_1-\frac{1}{2}x_3x_2+ \quad 4x_3^2.
\end{aligned}
$$

那么，f 的矩阵就是

$$
A=\begin{pmatrix} 1 & -1 & \frac{3}{2} \\ -1 & 0 & -\frac{1}{2} \\ \frac{3}{2} & -\frac{1}{2} & 4 \end{pmatrix}.
$$

这是一个对称矩阵. 此时 f 可表作

$$
f=(x_1 \quad x_2 \quad x_3)\begin{pmatrix} 1 & -1 & \frac{3}{2} \\ -1 & 0 & -\frac{1}{2} \\ \frac{3}{2} & -\frac{1}{2} & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=X'AX.
$$

例 2 给定二次型

$$
f=x_1x_3+x_2x_4.
$$

按上述办法写出它的矩阵为

$$
A=\begin{pmatrix} 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \end{pmatrix}.
$$

于是 f 可表作 $f=X'AX$.

如果给定数域 K 上一个 n 阶对称矩阵 $A=(a_{ij})$，那么，反过来以 a_{ij} 为系数即得一个二次型. 我们来证明：不同的对称矩阵对应于不同的二次型函数.

命题 2.1 给定数域 K 上两个二次型

$$f = X'AX = \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j \quad (a_{ij} = a_{ji}),$$

$$g = X'BX = \sum_{i=1}^{n}\sum_{j=1}^{n}b_{ij}x_ix_j \quad (b_{ij} = b_{ji}).$$

如果 $f\equiv g$（即 f,g 作为变元 $x_1,\cdots,x_n$ 的函数恒等），则 $A=B$.

证 设 V 内对称双线性函数 $f(\alpha,\beta)$，$g(\alpha,\beta)$ 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的矩阵分别为 A,B，则二次型函数 $Q_f(\alpha)$，$Q_g(\alpha)$ 在此组基下的解析表达式分别为二次型 f,g，故由 $f\equiv g$ 推知 $Q_f(\alpha)\equiv Q_g(\alpha)$，这又推出 $f(\alpha,\beta)\equiv g(\alpha,\beta)$，于是 $A=B$. ▌

现在我们有四个基本概念：对称双线性函数，二次型函数，二次型，K 上 n 阶对称矩阵. 它们在本质上是一样的. 我们可以用下面的图式把它们之间的关系表达出来：

$$f(\alpha,\beta)\xrightarrow{\beta=\alpha}Q_f(\alpha)\xrightarrow{\text{基 }\varepsilon_1,\cdots,\varepsilon_n}f=X'AX\Longrightarrow A\Longrightarrow f(\alpha,\beta),$$

就是说，给定 V 内对称双线性函数 $f(\alpha,\beta)$，令 $\beta=\alpha$ 即得二次型函数 $Q_f(\alpha)$，在 V 内取基 $\varepsilon_1,\cdots,\varepsilon_n$，得 $Q_f(\alpha)$ 的解析表达式 $X'AX$，它就是二次型 f，f 的系数组成 K 上对称矩阵 A，而 A 在基 $\varepsilon_1,\cdots,\varepsilon_n$ 下又唯一决定 V 内一对称双线性函数 $f(\alpha,\beta)$.

在 V 内另取一组基 $\eta_1,\eta_2,\cdots,\eta_n$，设

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T,$$

于是 $T=(t_{ij})$ 为两组基间的过渡矩阵. 又设

$$\alpha=y_1\eta_1+y_2\eta_2+\cdots+y_n\eta_n=(\eta_1,\eta_2,\cdots,\eta_n)Y,$$

那么，我们有熟知的坐标变换公式 $X=TY$，现在二次型函数 $Q_f(\alpha)$ 在基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的表达式是

$$Q_f(\alpha)=X'AX=(TY)'A(TY)=Y'(T'AT)Y,$$

这里 $T'AT$ 为 $f(\alpha,\beta)$ 在基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的矩阵.

换用数学分析的语言，我们给出如下定义：

定义 考察如下的 n 个变量的线性变数替换

$$\begin{cases} x_1 = t_{11}y_1 + t_{12}y_2 + \cdots + t_{1n}y_n, \\ x_2 = t_{21}y_1 + t_{22}y_2 + \cdots + t_{2n}y_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = t_{n1}y_1 + t_{n2}y_2 + \cdots + t_{nn}y_n. \end{cases} \tag{2}$$

如果其系数矩阵

$$T=\begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{pmatrix} \quad (t_{ij}\in K)$$

可逆，则(2)式称为**可逆线性变数替换**.

令

$$X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad Y=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

则(2)式可表示成矩阵形式

$$X=TY.$$

设 $S=(s_{ij})=T^{-1}$，则有 $Y=SX$. 具体写出就是

$$\begin{cases} y_1 = s_{11}x_1 + s_{12}x_2 + \cdots + s_{1n}x_n, \\ y_2 = s_{21}x_1 + s_{22}x_2 + \cdots + s_{2n}x_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n = s_{n1}x_1 + s_{n2}x_2 + \cdots + s_{nn}x_n. \end{cases} \tag{3}$$

(3)式称为(2)式的**逆变换**.

如果对二次型(1)做上述可逆线性变数替换，我们有

$$f=X'AX \xlongequal{X=TY} (TY)'A(TY)=Y'(T'AT)Y=g.$$

变换后得出 K 上另一个二次型 g(这时它的自变量用字母 $y_1,y_2,\cdots,y_n$ 表示. 从数学分析的观点看,函数中的自变量使用什么字母是无关紧要的). g 恰为 $Q_f(\alpha)$ 在新基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的解析表达式. g 的矩阵 $T'AT$ 为 $f(\alpha,\beta)$ 在基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的矩阵,与原矩阵 A 合同. 我们又有如下基本事实:

命题 2.2 给定 K 上两个二次型

$$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j \quad (a_{ij}=a_{ji}),$$

$$g=\sum_{i=1}^{n}\sum_{j=1}^{n}b_{ij}y_iy_j \quad (b_{ij}=b_{ji}).$$

它们的矩阵分别为 $A=(a_{ij}),B=(b_{ij})$. 则存在 K 上可逆线性变数替换 $X=TY$,使 f 变成 g 的充分必要条件是 B 与 A 合同:

$$B=T'AT.$$

证 若 f 经变换 $X=TY$ 化为 g,即

$$f=X'AX\xlongequal{X=TY}Y'(T'AT)Y=g=Y'BY,$$

那么,按命题 2.1,有

$$B=T'AT.$$

反之,若 $B=T'AT$ 且 $|T|\neq 0$,作变数替换 $X=TY$,有

$$f=X'AX\xlongequal{X=TY}Y'(T'AT)Y=Y'BY=g. \blacksquare$$

给定数域 K 上两个二次型 f,g,若 f 可经可逆线性变数替换化为 g,则称 f 与 g 等价,记作 $f\sim g$. 命题 2.2 说明,f 与 g 等价的充分必要条件是它们的矩阵合同. 合同关系是数域 K 上 n 阶对称矩阵集合中的一个等价关系(与对称矩阵 A 合同的矩阵 B 仍为对称矩阵:$B=T'AT$,则 $B'=(T'AT)'=T'A'T=T'AT=B$). 因此,二次型的合同关系也是数域 K 上全体二次型所成的集合内的一个等价关系,于是二次型按此关系划分为互不相交的等价类. 二次型理论的一个基本问题,就是要从每个等价类中挑选出一个最简单的二次型来作为该等价类的代表.

形如

$$d_1z_1^2+d_2z_2^2+\cdots+d_nz_n^2=Z'DZ$$

的二次型称为**标准形**，其矩阵为对角形

$$D=\begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}.$$

显然，这种二次型是较简单的.

根据§1中的定理1的推论，任意一个对称方阵都合同于一个对角矩阵. 从命题2.2可知，定理1可用二次型的语言叙述如下：

定理 1′ 给定数域 K 上的一个二次型

$$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j \quad (a_{ij}=a_{ji}),$$

则存在 K 上一个可逆方阵 T，使在线性变数替换 $X=TZ$ 下此二次型变为如下标准形

$$d_1z_1^2+d_2z_2^2+\cdots+d_nz_n^2.$$

综合§1和本节中的论述，可以把我们所得到的基本结果用三种不同的语言叙述出来.

(1) 用双线性函数的语言：数域 K 上 n 维线性空间 V 内任一对称双线性函数的矩阵都可对角化(即 V 内存在一组基，使该对称双线性函数在此组基下的矩阵成对角形).

(2) 用矩阵论的语言：数域 K 上任一对称的 n 阶方阵都合同于一个对角矩阵.

(3) 用二次型的语言：数域 K 上任一二次型都可经一个可逆线性变数替换化为标准形.

上述三种说法互相等价，只要证明其中之一，另两个就自动成立.

*二次型标准形的计算方法

我们现在介绍二次型的标准形的计算方法.分两种情况进行讨论.

1. 二次型(1)中有某个变量平方项的系数不为零,例如 $a_{11}\neq 0$.此时把二次型对 x_1 配方

$$
\begin{aligned}
f &= a_{11}x_1^2 + 2a_{12}x_1x_2 + \cdots + 2a_{1n}x_1x_n + \sum_{i=2}^{n}\sum_{j=2}^{n}a_{ij}x_ix_j \\
&= a_{11}\left(x_1^2 + 2x_1\cdot\frac{a_{12}}{a_{11}}x_2 + \cdots + 2x_1\cdot\frac{a_{1n}}{a_{11}}x_n\right) + \sum_{i=2}^{n}\sum_{j=2}^{n}a_{ij}x_ix_j \\
&= a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n\right)^2 + \sum_{i=2}^{n}\sum_{j=2}^{n}b_{ij}x_ix_j,
\end{aligned}
$$

其中 b_{ij}是由前面与 x_1 有关的项配方后出现的 $x_ix_j(i,j\geqslant 2)$项的系数和后面原有的 x_ix_j 项系数合并而成.

作变数替换

$$
\begin{cases}
y_1 = x_1 + \dfrac{a_{12}}{a_{11}}x_2 + \cdots + \dfrac{a_{1n}}{a_{11}}x_n, \\
y_2 = \qquad\qquad x_2, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
y_n = \qquad\qquad\qquad\qquad x_n.
\end{cases}
$$

反解为

$$
\begin{cases}
x_1 = y_1 - \dfrac{a_{12}}{a_{11}}y_2 - \cdots - \dfrac{a_{1n}}{a_{11}}y_n, \\
x_2 = \qquad\qquad y_2, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
x_n = \qquad\qquad\qquad\qquad y_n.
\end{cases}
$$

写成矩阵形式

$$X=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}1&-\frac{a_{12}}{a_{11}}&\cdots&-\frac{a_{1n}}{a_{11}}\\&1&&0\\&0&\ddots&\\&&&1\end{pmatrix}\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}.$$

不难看出,这实际上相当于定理 1 证明过程中所做的基变换. 经过上述变数替换后,二次型化作

$$a_{11}y_1^2+\sum_{i=2}^{n}\sum_{j=2}^{n}b_{ij}y_iy_j\quad(b_{ij}=b_{ji}).$$

然后再对上式右边的 $n-1$ 个变量 $y_2,y_3,\cdots,y_n$ 的二次型继续进行计算.

如果 $a_{11}=0$,而某个 $a_{ii}\neq0$,则对 x_i 配方.

2. 所有 $a_{ii}=0(i=1,2,\cdots,n)$,而有一个 $a_{ij}\neq0(i<j)$,则做变数替换

$$\begin{cases}x_i=y_i+y_j,\\x_j=y_i-y_j,\\x_k=y_k\quad(k\neq i,j).\end{cases}$$

这就可以把二次型化为第一种情况.

在做计算时,每一步都要把坐标变换矩阵记录下来,以便求得总的坐标变换矩阵. 下面举两个例子.

例 3 化二次型

$$f=2x_1x_2-6x_1x_3+2x_2x_3+x_3^2$$

为标准形.

解 因为 $a_{33}=1\neq0$,对 x_3 配方,得

$$\begin{aligned}f&=2x_1x_2+[2(x_2-3x_1)x_3+x_3^2]\\&=2x_1x_2-(x_2-3x_1)^2+[(-3x_1+x_2)+x_3]^2\\&=-9x_1^2+8x_1x_2-x_2^2+(-3x_1+x_2+x_3)^2.\end{aligned}$$

做变数替换

$$\begin{cases} y_1 = x_1, \\ y_2 = \qquad\quad x_2, \\ y_3 = -3x_1 + x_2 + x_3. \end{cases}$$

反解为
$$\begin{cases} x_1 = y_1, \\ x_2 = \qquad\quad y_2, \\ x_3 = 3y_1 - y_2 + y_3. \end{cases}$$

表成矩阵形式

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 3 & -1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = T_1 Y.$$

于是二次型化作

$$-9y_1^2 + 8y_1y_2 - y_2^2 + y_3^2.$$

y_3 已配好，下面把它保持不动. 现在 y_1^2 的系数不为零，对 y_1 配方，得

$$-9\left(y_1^2 - 2y_1 \cdot \frac{4}{9}y_2\right) - y_2^2 + y_3^2$$
$$= -9\left(y_1 - \frac{4}{9}y_2\right)^2 + \frac{7}{9}y_2^2 + y_3^2.$$

做变数替换，并表成矩阵形式

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & \frac{4}{9} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = T_2 Z.$$

于是二次型化作标准形

$$-9z_1^2 + \frac{7}{9}z_2^2 + z_3^2.$$

最后，求出总的变数替换矩阵 T. 因

$$X = T_1 Y = T_1(T_2 Z) = (T_1 T_2) Z = TZ,$$

故

$$T=T_1T_2=\begin{pmatrix}1&0&0\\0&1&0\\3&-1&1\end{pmatrix}\begin{pmatrix}1&\frac{4}{9}&0\\0&1&0\\0&0&1\end{pmatrix}$$

$$=\begin{pmatrix}1&\frac{4}{9}&0\\0&1&0\\3&\frac{1}{3}&1\end{pmatrix}.$$

在变换过程中得到的 T_1,T_2 均可逆，所以，它们的乘积 T 也是可逆的. 现在分别把原二次型的矩阵 A 和标准形的矩阵 D 写出来：

$$A=\begin{pmatrix}0&1&-3\\1&0&1\\-3&1&1\end{pmatrix},\quad D=\begin{pmatrix}-9&0&0\\0&\frac{4}{9}&0\\0&0&1\end{pmatrix}.$$

如果具体计算一下，就可以验证有合同关系：$T'AT=D$.

例 4 化二次型

$$f=2x_1x_2-6x_2x_3+2x_1x_3$$

为标准形.

解 现在所有平方项系数都为零，而 $a_{12}=1\neq 0$. 做变数替换

$$\begin{cases}x_1=y_1+y_2,\\x_2=y_1-y_2,\\x_3=\qquad\quad y_3.\end{cases}$$

写成矩阵形式

$$X=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}1&1&0\\1&-1&0\\0&0&1\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}=T_1Y.$$

二次型化作

$$2(y_1+y_2)(y_1-y_2)+2(y_1+y_2)y_3-6(y_1-y_2)y_3$$
$$=2y_1^2-2y_2^2-4y_1y_3+8y_2y_3.$$

y_1^2 的系数不为零,对 y_1 配方,得

$$2(y_1^2-2y_1y_3)-2y_2^2+8y_2y_3$$
$$=2(y_1-y_3)^2-2y_2^2-2y_3^2+8y_2y_3.$$

做变数替换,并写成矩阵形式

$$Y=\begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix}=T_2Z.$$

二次型化作

$$2z_1^2-2z_2^2+8z_2z_3-2z_3^2.$$

再对 z_2 配方,得

$$2z_1^2-2(z_2^2-4z_2z_3)-2z_3^2=2z_1^2-2(z_2-z_3)^2+6z_3^2.$$

做变数替换,并写成矩阵形式

$$Z=\begin{pmatrix}1&0&0\\0&1&2\\0&0&1\end{pmatrix}\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix}=T_3U.$$

二次型化作标准形

$$2u_1^2-2u_2^2+6u_3^2.$$

最后,求出总的坐标变换矩阵.因

$$X=T_1Y=T_1(T_2Z)=T_1T_2(T_3U)=(T_1T_2T_3)U,$$

故 $$T=T_1T_2T_3=\begin{pmatrix}1&1&0\\1&-1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&2\\0&0&1\end{pmatrix}$$
$$=\begin{pmatrix}1&1&3\\1&-1&-1\\0&0&1\end{pmatrix}.$$

如果写出原二次型矩阵 A 和标准形矩阵 D,有

$$A=\begin{pmatrix}0&1&1\\1&0&-3\\1&-3&0\end{pmatrix},\quad D=\begin{pmatrix}2&0&0\\0&-2&0\\0&0&6\end{pmatrix}.$$

不难验证有合同关系：$T'AT=D$.

在用变数替换化二次型成标准形时，必须注意所做的线性变数替换一定要是可逆的，即其系数矩阵行列式应当不为零. 不难证明，利用上面的办法做变换时，每一步的变换矩阵都是可逆的，所以，最后把它们连乘起来所得的总的变换矩阵 T 也一定是可逆的.

另外还应当注意，一个二次型的标准形并不是唯一的. 用不同的办法计算有可能得到不同的标准形. 关于这一点，我们在下一节中再做仔细研究.

习 题 二

1. 写出下列二次型的矩阵

(1) $f=-2x_1^2-x_2^2+x_1x_3-x_2x_3$；

(2) $f=-x_1x_3-2x_1x_4+x_3^2-5x_3x_4$；

(3) $f=2x_1^2-3x_2^2-4x_3^2-5x_4^2$；

(4) $f=-x_2^2-x_3^2+x_1x_4$.

2. 在 K^4 中给定如下对称双线性函数：若

$$\alpha=(x_1,x_2,x_3,x_4),\quad \beta=(y_1,y_2,y_3,y_4),$$

令

$$\begin{aligned}f(\alpha,\beta)=&-x_1y_1+2x_1y_2+2x_2y_1-3x_2y_2+x_2y_4\\&-2x_1y_4+x_4y_2-2x_4y_1+2x_4y_4.\end{aligned}$$

(1) 写出 $f(\alpha,\beta)$ 在基

$$\begin{aligned}\varepsilon_1&=(1,0,0,0),\\ \varepsilon_2&=(0,1,0,0),\\ \varepsilon_3&=(0,0,1,0),\\ \varepsilon_4&=(0,0,0,1)\end{aligned}$$

下的矩阵，并写出 $f(\alpha,\alpha)$ 在此组基下的解析表达式；

(2) 做基变换

$$(\eta_1,\eta_2,\eta_3,\eta_4)=(\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4)\begin{pmatrix}2 & 1 & \frac{4}{\sqrt{3}} & 0\\ 1 & 0 & \frac{3}{\sqrt{3}} & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & \frac{1}{\sqrt{3}} & 0\end{pmatrix},$$

求 $f(\alpha,\beta)$在基 $\eta_1,\eta_2,\eta_3,\eta_4$ 下的矩阵，并求 $f(\alpha,\alpha)$这组基下的解析表达式；

(3) 求可逆线性变数替换 $X=TY$，使二次型

$$f=-x_1^2+4x_1x_2-4x_1x_4-3x_2^2+2x_2x_4+2x_4^2$$

化成标准形.

3. 给定四个变量的二次型 f，试在 K^4 内找出对称双线性函数 $f(\alpha,\beta)$，使 $f(\alpha,\alpha)$在基

$$\varepsilon_1=(1,0,0,0),\quad \varepsilon_2=(0,1,0,0),$$
$$\varepsilon_3=(0,0,1,0),\quad \varepsilon_4=(0,0,0,1)$$

下的解析表达式为 f. 其中

(1) $f=x_1^2+x_2^2+x_3^2+x_4^2+2x_1x_2+2x_2x_3+2x_3x_4$；

(2) $f=x_1x_2+x_1x_3+x_1x_4+x_2x_3+x_2x_4+x_3x_4$.

4. 用可逆线性变数替换化下列二次型为标准形：

(1) $f=-4x_1x_2+2x_1x_3+2x_2x_3$；

(2) $f=x_1^2+2x_1x_2+2x_2^2+4x_2x_3+4x_3^2$；

(3) $f=x_1^2-3x_2^2-2x_1x_2+2x_1x_3-6x_2x_3$；

(4) $f=8x_1x_4+2x_3x_4+2x_2x_3+8x_2x_4$；

(5) $f=x_1x_2+x_1x_3+x_1x_4+x_2x_3+x_2x_4+x_3x_4$；

(6) $f=x_1^2+2x_2^2+x_4^2+4x_1x_2+4x_1x_3+2x_1x_4$
$+2x_2x_3+2x_2x_4+2x_3x_4$；

(7) $f=x_1^2+x_2^2+x_3^2+x_4^2+2x_1x_2+2x_2x_3+2x_3x_4$；

*(8) $f=x_1x_{2n}+x_2x_{2n-1}+\cdots+x_nx_{n+1}$；

*(9) $f=x_1x_2+x_2x_3+\cdots+x_{n-1}x_n$.

*5. 用可逆线性变数替换化下列二次型成标准形：

(1) $\sum_{i=1}^{n}x_i^2+\sum_{1\leqslant i<j\leqslant n}x_ix_j$;

(2) $\sum_{i=1}^{n}(x_i-\bar{x})^2$,其中 $\bar{x}=\dfrac{x_1+x_2+\cdots+x_n}{n}$.

6. 证明：秩等于 r 的对称矩阵可以表成 r 个秩等于 1 的对称矩阵之和.

*7. 给定二次型

$$f=\sum_{i=1}^{s}(a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n)^2,$$

其中

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{s1}&a_{s2}&\cdots&a_{sn}\end{pmatrix}$$

是一个实数矩阵. 证明 f 的秩等于 $r(A)$.

§3 实与复二次型的分类

上一节已经指出：我们可以把数域 K 上的二次型按等价关系划分为等价类. 现在要问：K 上的二次型究竟有多少个不同的等价类？这一节中，我们将要对 K 为复数域和实数域两种情况分别解决这个问题.

根据定理 1′，任一个二次型等价于一个标准形. 但二次型的标准形并不是唯一的，所以单靠标准形的概念还解决不了上面所提出的问题. 但标准形可以作为解决上述问题的出发点. 这就是说，只要讨论 $\mathscr{C}$ 或 $\mathscr{R}$ 上的二次型有多少种互不等价的标准形就可以了.

复二次型的分类

在复数域 $\mathscr{C}$ 上给定二次型

$$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j \quad (a_{ij}=a_{ji}).$$

设它在可逆线性变数替换 $X=TZ$ 下变为标准形

$$d_1z_1^2+d_2z_2^2+\cdots+d_nz_n^2.$$

根据§2中所作的阐述，这相当于在 $\mathscr{C}$ 上 n 维线性空间 V 内做一个基变换

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T,$$

使对称双线性函数 $f(\alpha,\beta)$ 在新基下的矩阵成对角形，即

$$f(\eta_i,\eta_j)(\eta_i,\eta_j)=d_i\delta_{ij},$$

其中 δ_{ij} 为克朗涅克记号. 设 $d_1,d_2,\cdots,d_n$ 中有 r 个不为零. 只要把 $\eta_1,\eta_2,\cdots,\eta_n$ 的次序重新排列一下，就可以使不为零的 d_i 排在前面，而后面 $n-r$ 个 d_i 全为零. 因此，不妨设 f 的标准形为

$$d_1z_1^2+d_2z_2^2+\cdots+d_rz_r^2 \quad (d_i\neq0,i=1,2,\cdots,r),$$

f 的矩阵为 $A=(a_{ij})$，有

$$T'AT=D=\begin{pmatrix} d_1 &&&&&& \\ & d_2 &&&&& \\ && \ddots &&&& \\ &&& d_r &&& \\ &&&& 0 && \\ &&&&& \ddots & \\ &&&&&& 0 \end{pmatrix}.$$

因为 T 可逆，由第二章命题2.3，$r(D)=r(A)$. 故 D 中主对角线上非零元素个数 $r=r(D)=r(A)=f$ 的秩.

因为在复数域内任意一个数都可以开平方，所以可以对上述标准形再做如下可逆线性变数替换：

$$\begin{cases} u_1 = \sqrt{d_1} z_1, \\ u_2 = \qquad \sqrt{d_2} z_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ u_r = \qquad\qquad \sqrt{d_r} z_r, \\ u_{r+1} = \qquad\qquad\qquad z_{r+1}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ u_n = \qquad\qquad\qquad\qquad z_n \end{cases}$$

(其中$\sqrt{d_i}$为d_i的任一平方根). 写成矩阵形式,就是

$$U = \begin{pmatrix} u_1 \\ \vdots \\ u_r \\ u_{r+1} \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} \sqrt{d_1} & & & & & \\ & \ddots & & & & \\ & & \sqrt{d_r} & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ \vdots \\ z_r \\ z_{r+1} \\ \vdots \\ z_n \end{pmatrix}.$$

于是f变作

$$u_1^2 + u_2^2 + \cdots + u_r^2.$$

定理 2 复数域$\mathscr{C}$上任一个二次型f都等价于如下一个二次型:

$$u_1^2 + u_2^2 + \cdots + u_r^2,$$

其中r等于二次型f的秩,这称为该二次型的**规范形**. 规范形是唯一的.

规范形由二次型的秩唯一决定,所以它是唯一的. 因为在证明中用到对某些数做开方运算,所以这个结论对一般数域不成立(一般数域上一个数开方后可能不再属该数域).

秩不同的二次型显然不等价. 定理 2 指出,在复数域内秩相同的二次型有相同的规范形,因而互相等价. 二次型的秩可为 0,1,2,$\cdots$,n,共$n+1$种可能,所以复数域$\mathscr{C}$上的二次型一共有$n+1$个不同的等价类.

实二次型的分类

在实数域 $\mathscr{R}$ 上给定二次型

$$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j \quad (a_{ij}=a_{ji}).$$

设 f 的秩为 r，那么，按上一段开头所指出的，存在 $\mathscr{R}$ 上可逆线性变数替换 $X=TZ$，使 f 化为标准形

$$d_1z_1^2+d_2z_2^2+\cdots+d_rz_r^2.$$

其中 $d_1,d_2,\cdots,d_r$ 为非零实数. 按同样的道理，我们不妨设前 p 个：$d_1,d_2,\cdots,d_p$ 为正数，而余下的 $r-p$ 个：$d_{p+1}\cdots,d_r$ 为负数. 因为在 $\mathscr{R}$ 内任何正数均可开平方，故可做 $\mathscr{R}$ 内可逆线性变数替换

$$\begin{cases} u_1=\sqrt{d_1}z_1, \\ \cdots\cdots\cdots\cdots \\ u_p=\qquad\sqrt{d_p}z_p, \\ u_{p+1}=\qquad\qquad\sqrt{-d_{p+1}}z_{p+1}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ u_r=\qquad\qquad\qquad\qquad\sqrt{-d_r}z_r, \\ u_{r+1}=\qquad\qquad\qquad\qquad\qquad z_{r+1}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ u_n=\qquad\qquad\qquad\qquad\qquad\qquad z_n. \end{cases}$$

于是二次型化作

$$u_1^2+\cdots+u_p^2-u_{p+1}^2-\cdots-u_r^2,$$

其中 $0\leqslant p\leqslant r$. 于是我们有：

定理 3 实数域 $\mathscr{R}$ 上任一个二次型 f 都等价于如下一个二次型：

$$u_1^2+\cdots+u_p^2-u_{p+1}^2-\cdots-u_r^2.$$

它称为 f 的**规范形**. 规范形是唯一的.

证 现在只需证明规范形的唯一性. 规范形中的 r 等于 f 的秩，是唯一确定的，我们只需证明正平方项的个数 p 也是唯一确定的就可以了.

设 f 有两个规范形

$$u_1^2+\cdots+u_p^2-u_{p+1}^2-\cdots-u_r^2,$$

$$v_1^2+\cdots+v_q^2-v_{q+1}^2-\cdots-v_r^2.$$

若 $p\neq q$，不妨设 $p>q$. 二次型 f 对应于 $\mathscr{R}$ 上 n 维线性空间 V 内的对称双线性函数 $f(\alpha,\beta)$. 上述假设表明在 V 内存在两组基

$$\eta_1,\eta_2,\cdots,\eta_n;$$

$$\eta_1',\eta_2',\cdots,\eta_n',$$

使 $f(\alpha,\beta)$ 在这两组基下的矩阵分别为

$$(f(\eta_i,\eta_j))=\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}\quad (p\text{个}\ 1); \tag{1}$$

$$(f(\eta_i',\eta_j'))=\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}\quad (q\text{个}\ 1). \tag{2}$$

令

$$M=L(\eta_1,\cdots,\eta_p,\eta_{r+1},\cdots,\eta_n);$$

$$N=L(\eta'_{q+1},\cdots,\eta'_r).$$

则 $\dim M=p+(n-r)$，$\dim N=r-q$. 故

$$\dim M+\dim N=n+(p-q)>n.$$

由第四章§3的维数公式，有

$$\begin{aligned}\dim M\cap N&=\dim M+\dim N-\dim(M+N)\\&\geqslant\dim M+\dim N-n>0,\end{aligned}$$

故 $M\cap N\neq\{0\}$. 设 $\alpha\in M\cap N$，$\alpha\neq 0$，有

(i) $\alpha\in M$，则

$$\alpha=a_1\eta_1+\cdots+a_p\eta_p+a_{r+1}\eta_{r+1}+\cdots+a_n\eta_n.$$

于是由(1)式有

$$f(\alpha,\alpha)=a_1^2+\cdots+a_p^2\geqslant 0.$$

(ii) $\alpha\in N$，则

$$\alpha=b_{q+1}\eta'_{q+1}+\cdots+b_r\eta'_{r+1}.$$

于是由(2)式有

$$f(\alpha,\alpha)=-b_{q+1}^2-\cdots-b_r^2<0$$

(因 $\alpha\neq 0$，故 $b_{q+1},\cdots,b_r$ 不全为零).

综合(i)与(ii)得到矛盾的结果. 这说明 $p>q$ 不成立，故 $p\leqslant q$. 再由对称性知，$q\leqslant p$，于是 $q=p$. ▌

定理3通常称为**惯性定律**. p 称为实二次型 f 的**正惯性指数**，$r-p$ 称为 f 的**负惯性指数**，它们的差：

$$p-(r-p)=2p-r$$

称为 f 的**符号差**.

定理3说明：对实二次型，不但秩不相同时不等价，就是在秩相同时，如果正惯性指数不相同，也不等价(这是由于规范形的唯一性，正惯性指数不同的规范形互相不等价). 而如果秩与正惯性指数相同时，它们都等价于同一个规范形，因而彼此等价. 所以在实二次型里，每个等价类决定于两个非负整数：r，p，其中 $0\leqslant r\leqslant$

$n, 0 \leqslant p \leqslant r$. 这就完全解决了实二次型的分类问题.

习 题 三

1. 用可逆线性变数替换将下列复二次型化为规范形：

(1) $f=2x_1x_2-6x_2x_3+2x_1x_3$;

(2) $f=-5x_1^2+3x_3^2-2x_4^2$;

(3) $f=(1+\mathrm{i})x_1^2-(\sqrt{2}+2\mathrm{i})x_2^2-3\mathrm{i}x_3^2$;

(4) $f=(-1-\mathrm{i})x_1x_2+2\mathrm{i}x_2^2$.

2. 用可逆线性变数替换将下列实二次型化为规范形，并求其秩，正、负惯性指数和符号差：

(1) $f=2x_1x_2-6x_2x_3+2x_1x_3$;

(2) $f=-5x_1^2+3x_3^2-2x_4^2$;

(3) $f=-2(x_1+x_2)^2+3(x_1-x_2)^2$;

(4) $f=-(x_1-2x_2+3x_3-x_4)^2+(2x_1-x_3+3x_4)^2$
$+(-2x_1+4x_2-6x_3+2x_4)^2$.

*3. 证明：一个实二次型可以分解成两个实系数的一次齐次多项式的乘积的充分必要条件是：它的秩等于 2，而符号差为零，或其秩为 1.

4. 设

$$f=X'AX=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j \quad (a_{ij}=a_{ji})$$

是一个实二次型. 若存在 n 维实向量 X_1 与 X_2，使

$$X_1'AX_1>0, \quad X_2'AX_2<0,$$

证明：必存在 n 维实向量 $X_0\neq 0$，使 $X_0'AX_0=0$.

*5. 设

$$f=l_1^2+l_2^2+\cdots+l_p^2-l_{p+1}^2-\cdots-l_{p+q}^2,$$

其中 $l_i(i=1,2,\cdots,p+q)$ 是 $x_1,x_2,\cdots,x_n$ 的实系数的一次齐次函数. 证明 f 的正惯性指数 $\leqslant p$，负惯性指数 $\geqslant q$.

6. 设 V 是实数域上的 n 维线性空间，$f(\alpha,\beta)$ 是 V 内一个对

称双线性函数，A 为 $f(\alpha,\beta)$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵. 若实二次型 $X'AX$ 的负惯性指数 $q=0$，证明

$$M=\{\alpha\in V \mid f(\alpha,\alpha)=0\}$$

是 V 的一个子空间，并求 $\dim M$.

§4 正定二次型

在这一节里，我们研究实二次型中秩 r 和正惯性指数 p 都等于 n 的等价类. 这个等价类在理论上具有特殊的重要性. 例如在下一章中，我们将要利用它对实数域上的线性空间作深入一步的研究.

定义 实数域 $\mathscr{R}$ 上的一个二次型

$$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j \quad (a_{ij}=a_{ji}), \tag{1}$$

如果它的秩 r 和正惯性指数都等于变量个数 n，则称为一个**正定二次型**. 正定二次型的矩阵称为**正定矩阵**.

为着下面讨论正定二次型的需要，我们重申一下 §2 中所作的一个约定：二次型(1)的矩阵 $A=(a_{ij})$ 看作是 $\mathscr{R}$ 上一个 n 维线性空间 V 内一个对称双线性函数 $f(\alpha,\beta)$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵，而二次型(1)本身则看作是 $f(\alpha,\alpha)$ 在这组基下的解析表达式.

命题 4.1 对实数域 $\mathscr{R}$ 上的二次型(1)下列命题等价：

(i) 二次型(1)是正定的；

(ii) A 合同于单位矩阵 E，亦即 $f(\alpha,\beta)$ 在某一组基下的矩阵为 E；

(iii) 对一切 $\alpha\in V$，$f(\alpha,\alpha)\geqslant 0$，且 $f(\alpha,\alpha)=0$ 的充分必要条件是 $\alpha=0$.

证 采用轮转办法来证明之.

(i)⇒(ii)：若(1)正定，则它的规范形为

$$u_1^2+u_2^2+\cdots+u_n^2=U'EU.$$

于是存在 $\mathscr{R}$ 上的可逆矩阵 T，使(1)在变数替换 $X=TU$ 下变为上述规范形. 由命题 2.2，有 $T'AT=E$.

(i)⇒(ii)：若存在可逆矩阵 T，使 $T'AT=E$，则在 V 内做基变换

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T,$$

$f(\alpha,\beta)$在新基下的矩阵为 E. 若令

$$\alpha=u_1\eta_1+u_2\eta_2+\cdots+u_n\eta_n,$$

则

$$f(\alpha,\alpha)=u_1^2+u_2^2+\cdots+u_n^2.$$

显然此时有 $f(\alpha,\alpha)\geqslant 0$，且 $f(\alpha,\alpha)=0$ 的充分必要条件是：

$$u_1=u_2=\cdots=u_n=0,\quad \text{即}\quad \alpha=0.$$

(iii)⇒(i)：在 V 内做基变换

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T,$$

使 $f(\alpha,\alpha)$在新基下的解析表达式变为二次型(1)的规范形

$$f(\alpha,\alpha)=u_1^2+\cdots+u_p^2-u_{p+1}^2-\cdots-u_r^2, \tag{2}$$

其中

$$\alpha=u_1\eta_1+u_2\eta_2+\cdots+u_n\eta_n.$$

如果 $r<n$，则令 $\alpha_0=\eta_n$ 代入(2)，有 $f(\alpha_0,\alpha_0)=0$，而 $\alpha_0\neq 0$，与假设矛盾，故 $r=n$. 而如果 $p<r=n$，则以同样的 $\alpha_0=\eta_n\neq 0$ 代入(2)，得 $f(\alpha_0,\alpha_0)=-1<0$，亦与假设矛盾，故必 $p=r=n$. 即二次型正定. ▌

因为

$$f(\alpha,\alpha)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j,$$

其中 $\alpha=x_1\varepsilon_1+x_1\varepsilon_2+\cdots+x_n\varepsilon_n,$

从命题 4.1 的(iii)可知，f 正定的充分必要条件是

$$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\geqslant 0,$$

且等号成立的充分必要条件是 $x_1=x_2=\cdots=x_n=0$.

从命题 4.1 还可以得到如下的推论：

推论 如 A 是正定矩阵，则 $|A|>0$.

证 由命题 4.1 的(ii)知，A 合同于单位矩阵 E. 于是有实可逆矩阵 T，使 $A=T'ET=T'T$. 那么，有

$$|A|=|T'T|=|T'||T|=|T|^2>0. \blacksquare$$

命题 4.1 是二次型正定(或实对称矩阵正定)的几种等价的说法，其中每一种都可以用作定义. 下面再给出实二次型正定的一个重要判别法.

定义 设 A 是一个实对称矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

称如下 n 个行列式

$$|a_{11}|, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \cdots, \begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix}, \cdots, \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

为 A 的**顺序主子式**.

定理 4 给定 $\mathscr{R}$ 上二次型

$$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j \quad (a_{ij}=a_{ji}),$$

它的矩阵记为 $A=(a_{ij})$，则 f 正定的充分必要条件是，A 的所有顺序主子式都大于零，即

$$|a_{11}|>0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}>0, \quad \cdots,$$

$$\begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix}>0, \quad \cdots, \quad \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}>0.$$

证 (i) 必要性. 考虑 V 的子空间

$$M_r=L(\varepsilon_1,\cdots,\varepsilon_r)\quad(r=1,2,\cdots,n),$$

f 正定,则对任一 $\alpha\in M_r$,有 $f(\alpha,\alpha)\geqslant 0$,且 $f(\alpha,\alpha)=0$ 的充分必要条件是 $\alpha=0$. 但

$$\alpha=x_1\varepsilon_1+\cdots+x_r\varepsilon_r,$$

故

$$f(\alpha,\alpha)=\sum_{i=1}^{r}\sum_{j=1}^{r}a_{ij}x_ix_j\quad(a_{ij}=a_{ji}).$$

上式是 r 个变量 $x_1,x_2,\cdots,x_r$ 的正定二次型,由命题 4.1 的推论知,它的矩阵的行列式大于零,即

$$\begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix}>0\quad(r=1,2,\cdots,n).$$

(ii) 充分性. 对 n 作数学归纳法.

$n=1$ 时命题显然成立. 设命题对 $n-1$ 个变量的二次型成立,证明它对 n 个变量的二次型也成立.

命

$$M_{n-1}=L(\varepsilon_1,\cdots,\varepsilon_{n-1}).$$

$\alpha\in M_{n-1}$时,有

$$\alpha=x_1\varepsilon_1+\cdots+x_{n-1}\varepsilon_{n-1},$$

$$f(\alpha,\alpha)=\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}a_{ij}x_ix_j\quad(a_{ij}=a_{ji}).$$

上式是 $n-1$ 个变量的二次型,它的矩阵的 $n-1$ 个顺序主子式恰为 A 的前 $n-1$ 个顺序主子式,全为正,按归纳假设,这是一个正定二次型. 于是在 M_{n-1}内存在一组基

$$\eta_1,\eta_2,\cdots,\eta_{n-1},$$

使 $f(\alpha,\beta)$(看作 M_{n-1}内的一个对称双线性函数)在这组基下的矩阵变为单位矩阵,即

$$f(\eta_i,\eta_j)=\delta_{ij}\quad(i,j=1,2,\cdots,n-1).\tag{3}$$

把 M_{n-1}的这组基扩充为 V 的一组基

$$\eta_1,\eta_2,\cdots,\eta_{n-1},\xi. \tag{4}$$

令

$$\eta'_n=-f(\eta_1,\xi)\eta_1-f(\eta_2,\xi)\eta_2-\cdots-f(\eta_{n-1},\xi)\eta_{n-1}+\xi.$$

向量组

$$\eta_1,\cdots,\eta_{n-1},\eta'_n \tag{5}$$

与向量组(4)等价，故它也是V的一组基，且有(利用(3)式)

$$\begin{aligned} f(\eta_i,\eta'_n)&=-f(\eta_i,\xi)f(\eta_i,\eta_i)+f(\eta_i,\xi)\\ &=-f(\eta_i,\xi)+f(\eta_i,\xi)=0, \end{aligned} \tag{6}$$

其中$i=1,2,\cdots,n-1$. 由此知$f(\alpha,\beta)$在基(5)下的矩阵为

$$B=\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & f(\eta'_n,\eta'_n) \end{pmatrix}.$$

B与A是$f(\alpha,\beta)$在不同基下的矩阵，故$B=T'_1AT_1$，$|T_1|\neq0$. 于是有

$$f(\eta'_n,\eta'_n)=|B|=|T'_1AT_1|=|T_1|^2\cdot|A|>0.$$

令

$$\eta_n=\frac{1}{\sqrt{f(\eta'_n,\eta'_n)}}\eta'_n, \tag{7}$$

则$f(\alpha,\beta)$在基$\eta_1,\eta_2,\cdots,\eta_n$下的矩阵为

$$(f(\eta_i,\eta_j))=(\delta_{ij})=E$$

(其中利用了(3)，(6)，(7)式). 由命题4.1的(ii)知f正定. ▌

最后我们再介绍几个概念.

1. 如果实二次型f的正惯性指数p等于秩r，则f称为**半正定二次型**. 这种二次型的规范形为

$$u_1^2+u_2^2+\cdots+u_r^2.$$

显然，此时f对应的对称双线性函数$f(\alpha,\beta)$满足

$$f(\alpha,\alpha)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\geqslant 0,$$

其中

$$\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n.$$

2. 如果实二次型 f 的负惯性指数和秩 r 都等于 n，则 f 称为**负定二次型**. 这种二次型的规范形为

$$-u_1^2-u_2^2-\cdots-u_n^2.$$

此时有

$$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\leqslant 0,$$

等号成立的充分必要条件是 $x_1=x_2=\cdots=x_n=0$.

3. 如果实二次型 f 的负惯性指数等于秩 r，则 f 称为**半负定二次型**. 这种二次型的规范形为

$$-u_1^2-u_2^2-\cdots-u_r^2.$$

显然，此时有 $$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\leqslant 0.$$

4. 除了正定、半正定，负定、半负定之外的其它实二次型统称为**不定二次型**.

习　题　四

1. 判断下列二次型是否正定：

(1) $f=99x_1^2-12x_1x_2+48x_1x_3+130x_2^2-60x_2x_3+71x_3^2$；

(2) $f=10x_1^2+8x_1x_2+24x_1x_3+2x_2^2-28x_2x_3+x_3^2$；

*(3) $f=\sum_{i=1}^{n}x_i^2+\sum_{1\leqslant i<j\leqslant n}x_ix_j$；

*(4) $f=\sum_{i=1}^{n}x_i^2+\sum_{i=1}^{n-1}x_ix_{i+1}$.

2. t 取什么值时，下列二次型是正定的？

(1) $f=x_1^2+x_2^2+5x_3^2+2tx_1x_2-2x_1x_3+4x_2x_3$；

(2) $f=x_1^2+4x_2^2+x_3^2+2tx_1x_2+10x_1x_3+6x_2x_3$.

3. 证明:如果 A 是正定矩阵,那么 A 的主子式全大于零(主子式是指行角标和列角标相同的子式,即 A 的第 $i_1,i_2,\cdots,i_r$ 行,第 $i_1,i_2,\cdots,i_r$ 列交叉点处的元素所成的 r 阶子式).

4. 设 A 是实对称矩阵,证明:当实数 t 充分大之后,$tE+A$ 是正定矩阵.

5. 证明:如果 A 是正定矩阵,那么 A^{-1} 也是正定矩阵.

6. 设 A 是实的 n 阶对称矩阵,证明:存在一个正实数 c,使对任一实 n 维向量 X(表成列形式)都有

$$|X'AX|\leqslant cX'X.$$

7. 设 A 为 n 阶实对称矩阵,$|A|<0$. 证明:存在实的 n 维向量 X,使 $X'AX<0$.

8. 如果 A,B 都是正定矩阵,证明 $A+B$ 也是正定矩阵.

9. 证明:$n\sum_{i=1}^{n}x_i^2-\left(\sum_{i=1}^{n}x_i\right)^2$ 是半正定的.

10. 证明:

(1) 如果 $f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\quad(a_{ij}=a_{ji})$ 是正定二次型,那么

$$g(y_1,y_2,\cdots,y_n)=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & y_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & y_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & y_n \\ y_1 & y_2 & \cdots & y_n & 0 \end{vmatrix}$$

是负定二次型.

(2) 如果 A 是正定矩阵,那么

$$|A|\leqslant a_{nn}\cdot P_{n-1},$$

其中 P_{n-1} 是 A 的 $n-1$ 阶顺序主子式.

(3) 如果 A 是正定矩阵,那么

$$|A|\leqslant a_{11}a_{22}\cdots a_{nn}.$$

(4) 如果 $T=(t_{ij})$ 是 n 阶实可逆矩阵，那么

$$|T|^2\leqslant\prod_{i=1}^{n}(t_{1i}^2+t_{2i}^2+\cdots+t_{ni}^2).$$

11. 证明：实对称矩阵 A 是半正定的充分必要条件是：A 的一切主子式都 $\geqslant 0$.

第七章　欧几里得空间

在第四章中，我们引进了数域 K 上的线性空间的概念. 随后，在第五章和第六章中，我们研究了线性空间中的线性变换，以及线性函数、双线性函数、二次型函数，获得了许多重要的结果. 在做了这一系列工作之后，我们再回顾一下线性空间的概念本身，就会注意到，我们对线性空间概念本身的要求却是非常少的：仅仅要求它有一个加法运算和一个数乘运算，以及这两种运算所必须满足的八条公理. 除此之外，就没有其它的. 就在这样简单的基础之上，我们已经从理论上得到了相当丰富的成果. 但在另一方面，我们还必须注意到，我们关于线性空间的上述概念，是从三维几何空间抽象而来的，在这个抽象化的过程中，我们抛弃了三维几何空间向量的许多重要的几何性质. 例如，三维几何空间中向量的长度、夹角等与度量有关的性质，在我们前面所给出的线性空间概念中就没有得到任何反映. 可是，向量的长度与夹角这类概念不但在理论上，而且在实际工作中都是有重要意义的. 这就启发我们：应当进一步充实线性空间的概念，把与三维几何空间中向量的长度和夹角相类似的概念添加到一般的线性空间中来.

在空间解析几何中，我们已经知道，向量的与长度、夹角有关的度量性质完全取决于向量的内积(又称点乘或数量积). 如果我们能把内积的概念推广到一般的线性空间中去，那么，一般的线性空间也就会具有某些度量性质，这样，我们关于线性空间的概念也就大大地丰富了.

本章将集中精力研究在实数域上的线性空间中引进内积的问题，下一章再讨论复数域上的线性空间中的相应问题.

§1 欧几里得空间的定义和基本性质

欧几里得空间的定义

定义 设 V 是实数域 $\mathscr{R}$ 上的线性空间. 如果 V 内任意两个向量 α,β 都按某一法则对应于 $\mathscr{R}$ 内一个唯一确定的数,记作 (α,β),且满足:

(i) 对任意 $k_1,k_2\in\mathscr{R}$ 和任意 $\alpha_1,\alpha_2,\beta\in V$,有

$$(k_1\alpha_1+k_2\alpha_2,\beta),=k_1(\alpha_1,\beta)+k_2(\alpha_2,\beta);$$

(ii) 对任意 $\alpha,\beta\in V$,有

$$(\alpha,\beta)=(\beta,\alpha);$$

(iii) 对任意 $\alpha\in V$,有 $(\alpha,\alpha)\geqslant 0$,且 $(\alpha,\alpha)=0$ 的充分必要条件是 $\alpha=0$.

则称 (α,β) 为向量 α,β 的**内积**. 定义了内积的实数域上线性空间称为**欧几里得空间**,简称**欧氏空间**.

从性质(i)和(ii)可知,对任意 $l_1,l_2\in\mathscr{R}$ 和任意 $\alpha,\beta_1,\beta_2\in V$,有

$$\begin{aligned}(\alpha,l_1\beta_1+l_2\beta_2)&=(l_1\beta_1+l_2\beta_2,\alpha)\\&=l_1(\beta_1,\alpha)+l_2(\beta_2,\alpha)\\&=l_1(\alpha,\beta_1)+l_2(\alpha,\beta_2).\end{aligned}$$

把这性质和(i),(ii)结合起来就可以看出,(α,β) 实际上是 V 内一个对称双线性函数. 如果 V 是有限维的线性空间,那么性质(iii)表明 (α,α) 是一个正定二次型函数,即 (α,β) 在 V 的任一组基下的矩阵都是正定矩阵. 反过来,如果在 V 内给定一个对称双线性函数 $f(\alpha,\beta)$,且 $f(\alpha,\alpha)$ 是一个正定二次型函数,则只要把 V 内两个向量的内积定义为

$$(\alpha,\beta)=f(\alpha,\beta),$$

那么,V 关于这个内积成一欧氏空间. 由此可知:$\mathscr{R}$ 上有限维线性

空间的内积概念和正定二次型概念之间有密切的关系.

有了内积的定义,我们就可以利用它给出欧氏空间内向量的长度和夹角的概念.

一、向量的长度

对任意 $\alpha\in V$,定义

$$|\alpha|=\sqrt{(\alpha,\alpha)},$$

称为 α 的**长度**或**模**. 从内积的性质(iii)可知,$|\alpha|=0$ 的充分必要条件是 $\alpha=0$. $|\alpha|=1$ 时,α 称为**单位向量**.

对任一 $k\in\mathscr{R}$,有

$$|k\alpha|=\sqrt{(k\alpha,k\alpha)}=|k|\cdot|\alpha|.$$

由此知,当 $\alpha\neq0$ 时,$\left|\frac{1}{|\alpha|}\alpha\right|=\frac{1}{|\alpha|}\cdot|\alpha|=1$,即 $\frac{1}{|\alpha|}\alpha$ 为一单位向量,我们称它为 α 的**单位化**.

二、向量的夹角

为了定义一般欧氏空间 V 内向量夹角的概念,我们需要如下的命题:

命题 1.1 对欧氏空间 V 内任意两个向量 α,β,有

$$|(\alpha,\beta)|\leqslant|\alpha|\cdot|\beta|.$$

等号成立的充分必要条件是:α 与 β 线性相关.

证 $\alpha=0$ 或 $\beta=0$ 时命题显然正确. 现设 $\alpha\neq0,\beta\neq0$. 令 $\gamma=t\alpha+\beta$,则

$$\begin{aligned}0\leqslant(\gamma,\gamma)&=(t\alpha+\beta,t\alpha+\beta)\\&=(\alpha,\alpha)t^2+2(\alpha,\beta)t+(\beta,\beta).\end{aligned}$$

上式右端是 t 的二次多项式,其值恒$\geqslant0$,故它没有相异的实根(否则,因 t^2 项系数$(\alpha,\alpha)>0$,它在两实根之间函数值为负). 因而,其判别式

$$[2(\alpha,\beta)]^2-4(\alpha,\alpha)(\beta,\beta)\leqslant0,$$

故
$$|(\alpha,\beta)|\leqslant|\alpha|\cdot|\beta|.$$

显然等号成立(即判别式等于零)的充分必要条件是:上述二次多

项式有实根（二重根）$t=k$. 这等价于

$$(k\alpha+\beta,k\alpha+\beta)=0.$$

由内积的性质(iii)知，后者等介于 $k\alpha+\beta=0$. ▌

命题 1.1 称为**哥西-布尼雅可夫斯基**(Cauchy-Буняковский)**不等式**.

现在可以给出欧氏空间 V 内向量夹角的定义. 对 V 内任意两个非零向量 α,β，由命题 1.1，$\left|\frac{(\alpha,\beta)}{|\alpha|\cdot|\beta|}\right|\leqslant 1$，所以可以定义

$$<\alpha,\beta>=\arccos\frac{(\alpha,\beta)}{|\alpha|\cdot|\beta|},$$

称之为 α 与 β 的**夹角**. 注意这样定义的两向量的夹角总介于 0 与 π 之间. 零向量与其它向量的夹角认为是不确定的.

如果 $(\alpha,\beta)=0$，则称 α 与 β **正交**，记作 $\alpha\perp\beta$. 当 $\alpha\neq 0,\beta\neq 0$ 时，这与 $<\alpha,\beta>=\frac{\pi}{2}$ 等价. 显然，零向量与任意向量正交.

下面举几个例子.

例 1 考虑实数域上 n 维向量空间 $\mathscr{R}^n$，对

$$\alpha=(a_1,a_2\cdots,a_n);\quad \beta=(b_1,b_2,\cdots,b_n),$$

定义

$$(\alpha,\beta)=a_1b_1+a_2b_2+\cdots+a_nb_n.$$

显然，二元函数 (α,β) 满足内积定义中的条件(i)～(iii)，于是 $\mathscr{R}^n$ 关于这个内积成一欧氏空间. 我们约定：今后凡称 $\mathscr{R}^n$ 为欧氏空间时，其内积都是按上述法则定义的（除了有特别声明的情况而外）.

我们知道 $\mathscr{R}^n$ 有一组基

$$\varepsilon_1=(1,0,\cdots,0),$$
$$\varepsilon_2=(0,1,\cdots,0),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\varepsilon_n=(0,0,\cdots,1).$$

显然有 $|\varepsilon_i|=1(i=1,2\cdots,n)$，即 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 都是单位向量. 另一方面，不难算出 $(\varepsilon_i,\varepsilon_j)=0(i\neq j)$. 故这 n 个向量两两正交. 上面两条性质可简记为

$$(\varepsilon_i,\varepsilon_j)=\delta_{ij}.$$

在欧氏空间 $\mathscr{R}^n$ 内，向量的长度和夹角可分别用公式表示如下：

$$|\alpha|=\sqrt{a_1^2+a_2^2+\cdots+a_n^2};$$

$$<\alpha,\beta>=\arccos\frac{a_1b_1+a_2b_2+\cdots+a_nb_n}{\sqrt{a_1^2+a_2^2+\cdots+a_n^2}\cdot\sqrt{b_1^2+b_2^2+\cdots+b_n^2}}.$$

而哥西-布尼雅可夫斯基不等式可具体写成

$$|a_1b_1+a_2b_2+\cdots+a_nb_n|$$

$$\leqslant\sqrt{a_1^2+a_2^2+\cdots+a_n^2}\cdot\sqrt{b_1^2+b_2^2+\cdots+b_n^2}.$$

例 2 考虑闭区间 $[a,b]$ 上全体实连续函数所组成的实数域上线性空间 $C[a,b]$. 对任意 $f,g\in C[a,b]$，定义

$$(f,g)=\int_a^b f(x)g(x)\mathrm{d}x.$$

二元函数 (f,g) 显然满足内积定义中的条件(i)与(ii). 另一方面，显然有 $(f,f)\geqslant 0$. 而当

$$(f,f)=\int_a^b f^2(x)\mathrm{d}x=0$$

时，由定积分的知识知，在 $[a,b]$ 内 $f(x)\equiv 0$. 即 $f(x)$ 为区间 $[a,b]$ 上的零函数，从而是 $C[a,b]$ 中的零向量. 因而内积的三个条件均满足. 于是 $C[a,b]$ 关于这个内积成一欧氏空间. 在这欧氏空间内，哥西-布尼雅可夫斯基不等式可具体写成

$$\left|\int_a^b f(x)g(x)\mathrm{d}x\right|\leqslant\sqrt{\int_a^b f^2(x)\mathrm{d}x}\cdot\sqrt{\int_a^b g^2(x)\mathrm{d}x}.$$

例 3 考察 $\mathscr{R}[X]_n$. 用两种方式在这个线性空间内定义内积：

1. 对任意 $f,g\in\mathscr{R}[X]_n$，定义

$$(f,g)=\int_0^1 f(x)g(x)\mathrm{d}x.$$

内积条件(i)，(ii)显然满足，且 $(f,f)\geqslant 0$. 当

$$(f,f)=\int_0^2 f^2(x)\mathrm{d}x=0$$

时，在[0,1]内 $f(X)\equiv 0$. $f(X)$是次数$<n$ 的实系数多项式，当它不恒等于零时最多有 $n-1$ 个根，故必定在$(-\infty,\infty)$内$f(X)\equiv 0$. 于是 $\mathscr{R}[X]_n$ 关于上述内积成一欧氏空间.

2. 对任意 $f,g\in\mathscr{R}[X]_n$，定义

$$(f,g)=\sum_{k=1}^{n}f(k)g(k).$$

二元函数(f,g)显然满足内积条件(i)与(ii)，且$(f,f)\geqslant 0$. 而当

$$(f,f)=\sum_{k=1}^{n}f^2(k)=0$$

时，有 $f(1)=f(2)=\cdots=f(n)=0$. 即 $f(X)$有 n 个不同实根，而 $f(X)$的次数$\leqslant n-1$，故 $f(X)\equiv 0$. 于是 $\mathscr{R}[X]_n$ 关于这个新内积也成为一个欧氏空间.

对于实数域上的同一个线性空间 V，当我们用不同的办法来定义它的内积时，所得的欧氏空间认为是互不相同的.

有限维的欧氏空间

设 V 是一个 n 维欧氏空间，而

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$$

是它的一组基. 令

$$G=\begin{pmatrix}(\varepsilon_1,\varepsilon_1) & (\varepsilon_1,\varepsilon_2) & \cdots & (\varepsilon_1,\varepsilon_n)\\ (\varepsilon_2,\varepsilon_1) & (\varepsilon_2,\varepsilon_2) & \cdots & (\varepsilon_2,\varepsilon_n)\\ \vdots & \vdots & & \vdots\\ (\varepsilon_n,\varepsilon_1) & (\varepsilon_n,\varepsilon_2) & \cdots & (\varepsilon_n,\varepsilon_n)\end{pmatrix},$$

称 G 为内积(α,β)在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的**度量矩阵**，它实际上就是对称双线性函数(α,β)在这组基下的矩阵. 因此，它必定是一个实对称矩阵. 而且有

1. 内积(α,β)在任一组基下的度量矩阵都是正定矩阵；

2. 如果(α,β)在另一组基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的度量矩阵为 $\overline{G}=((\eta_i,\eta_j))$，而

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T,$$

则 $\overline{G}=T'GT$，即内积在不同基下的度量矩阵互相合同；

3. 内积可用其度量矩阵

$$G=(g_{ij}),\ \ g_{ij}=(\varepsilon_i,\varepsilon_j)\ \ (i,j=1,2,\cdots,n)$$

表达如下：

$$(\alpha,\beta)=X'GY=\sum_{i=1}^{n}\sum_{j=1}^{n}g_{ij}x_iy_j,$$

其中

$$\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n,$$

$$\beta=y_1\varepsilon_1+y_2\varepsilon_2+\cdots+y_n\varepsilon_n.$$

上面三条都是把第六章中关于双线性函数所获得的一般结论应用于内积 (α,β) 这一特殊的双线性函数而得到的.

现在我们给出有限维欧氏空间 V 中的类似于三维几何空间中的直角坐标系的一个重要概念. 我们先证一个命题.

命题 1.2 设欧氏空间 V 内 s 个非零向量 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 两两正交，则它们线性无关.

证 若

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=0,$$

两边用 $\alpha_i(i=1,2,\cdots,s)$ 作内积，有

$$k_1(\alpha_1,\alpha_i)+k_2(\alpha_2,\alpha_i)+\cdots+k_s(\alpha_s,\alpha_i)=k_i(\alpha_i,\alpha_i)=0.$$

因 $\alpha_i\neq 0$，故 $(\alpha_i,\alpha_i)\neq 0$，即有 $k_i=0$. 这表明 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关. ▌

定义 n 维欧氏空间 V 中 n 个两两正交的单位向量

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$$

称为 V 的一组**标准正交基**.

由命题 1.2 知，标准正交基是 V 的一组基. 显然，V 内 n 个向量 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是一组标准正交基，等价于

$$(\varepsilon_i,\varepsilon_j)=\delta_{ij}\ \ (i,j=1,2,\cdots,n),$$

即等价于内积 (α,β) 在这组基下的度量矩阵是单位矩阵 E.

上面例 1 中已给出 $\mathscr{R}^n$ 中的一组标准正交基.

下面我们讨论有关标准正交基的几个问题.

一、标准正交基的存在性

设 V 是 n 维欧氏空间,在 V 内任取一组基

$$\xi_1,\xi_2,\cdots,\xi_n,$$

已知内积在这组基下的度量矩阵 G 是一个正定矩阵. 由第六章命题 4.1 知,G 合同于单位矩阵,即有实可逆矩阵 T,使 $T'GT=E$. 令

$$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\xi_1,\xi_2,\cdots,\xi_n)T,$$

则 (α,β) 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的度量矩阵为 E,从而 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组标准正交基. 这证明任一有限维欧氏空间都存在标准正交基.

二、两组标准正交基间的过渡矩阵

定义 设 $\mathscr{R}$ 上一个 n 阶方阵 T 满足

$$T'T=E,$$

亦即 $T'=T^{-1}$,则称 T 为**正交矩阵**.

显然,定义中的 $T'T=E$ 也可换成 $TT'=E$.

例如二维几何平面上两个直角坐标系的坐标变换矩阵

$$T=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

就满足 $T'T=E$,所以这是一个二阶正交矩阵. 对一般有限维欧氏空间,我们有与此相应的结论.

命题 1.3 在 n 维欧氏空间 V 内给定一组标准正交基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n,$$

令

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T,$$

则 $\eta_1,\eta_2,\cdots,\eta_n$ 是一组标准正交基的充分必要条件是:T 是一个正交矩阵.

证 (i) 必要性. 若 $\eta_1,\eta_2,\cdots,\eta_n$ 与 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 都是标准正交基,则 (α,β) 在这两组基下的度量矩阵都是 E,而且它们合同:

$T'ET=E$，即 $T'T=E$，于是 $T'=T^{-1}$，即 T 为正交矩阵.

(ii) 充分性. 若 T 是正交矩阵，则 T 可逆，于是 $\eta_1,\eta_2,\cdots,\eta_n$ 也是 V 的一组基. (α,β) 在这组基下的矩阵 G 与它在 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的度量矩阵 E 合同：$G=T'ET=T'T=E$. 故 $\eta_1,\eta_2,\cdots,\eta_n$ 是 V 的一组标准正交基. ▌

命题 1.3 给出了正交矩阵的一个等价定义：正交矩阵就是标准正交基之间的过渡矩阵. 下面再给出一个等价表述. 设

$$T=\begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{pmatrix},$$

如果我们把 T 的行向量组

$$\alpha_i=(t_{i1},t_{i2}\cdots,t_{in}) \quad (i=1,2,\cdots,n)$$

看作欧氏空间 $\mathscr{R}^n$ 中的一个向量组，则 $TT'=E$ 等价于

$$(\alpha_i,\alpha_j)=\delta_{ij} \quad (i,j=1,2,\cdots,n).$$

即 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 $\mathscr{R}^n$ 的一组标准正交基. 同样地，把正交矩阵 T 的列向量组看作 $\mathscr{R}^n$ 中的一个向量组，那么 $T'T=E$ 等价于这个列向量组构成 $\mathscr{R}^n$ 的一组标准正交基. 所以，我们有：

命题 1.4 $\mathscr{R}$ 上一个 n 阶方阵 T 是正交矩阵的充分必要条件是：T 的行或列向量组构成欧氏空间 $\mathscr{R}^n$ 的一组标准正交基.

如果用 T 的元素把它的正交性具体写出，就是

$$t_{i1}t_{j1}+t_{i2}t_{j2}+\cdots+t_{in}t_{jn}=\delta_{ij};$$

$$t_{1i}t_{1j}+t_{2i}t_{2j}+\cdots+t_{ni}t_{nj}=\delta_{ij}.$$

上面两组关系式中只要有一组成立（即 $TT'=E$ 或 $T'T=E$ 中有一式成立），则 T 就是正交矩阵，此时另一组关系式也随之成立.

三、标准正交基的求法

下面我们介绍具体寻求欧氏空间 V 的标准正交基的办法，这个方法通常称为**施密特**(Schmidt)**正交化方法**.

我们把问题提的更一般一些：给定 V 中一个线性无关的向量组

$$\alpha_1,\alpha_2,\cdots,\alpha_s. \qquad (\text{I})$$

要求作出一个新向量组

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_s, \qquad (\text{II})$$

满足如下两个条件：

(i) $L(\varepsilon_1,\cdots,\varepsilon_i)=L(\alpha_1,\cdots,\alpha_i)(i=1,2,\cdots,s)$；

(ii) $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_s$ 两两正交.

向量组(Ⅱ)可用如下办法给出

$$\varepsilon_1=\alpha_1,$$

$$\varepsilon_2=\alpha_2-\frac{(\alpha_2,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1,$$

$$\varepsilon_3=\alpha_3-\frac{(\alpha_3,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1-\frac{(\alpha_3,\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\varepsilon_2,$$

……………………………………

$$\varepsilon_{i+1}=\alpha_{i+1}-\frac{(\alpha_{i+1},\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1-\frac{(\alpha_{i+1},\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\varepsilon_2-\cdots-\frac{(\alpha_{i+1},\varepsilon_i)}{(\varepsilon_i,\varepsilon_i)}\varepsilon_i,$$

…………………………………………………………

$$\varepsilon_s=\alpha_s-\frac{(\alpha_s,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1-\frac{(\alpha_s,\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\varepsilon_2-\cdots-\frac{(\alpha_s,\varepsilon_{s-1})}{(\varepsilon_{s-1},\varepsilon_{s-1})}\varepsilon_{s-1}.$$

不难看出，上面构造出来的向量组 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_s$ 具有所要求的条件：

(i) 把上述等式右方带负号的项移到左端，即可看出向量组(Ⅰ)的前 i 个向量 $\alpha_1,\cdots,\alpha_i$ 可由向量组(Ⅱ)的前 i 个向量 $\varepsilon_1,\cdots,\varepsilon_i$ 线性表示. 反过来，因为 $\varepsilon_1=\alpha_1$，而

$$\varepsilon_2=\alpha_2-\frac{(\alpha_2,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1=\alpha_2-\frac{(\alpha_2,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\alpha_1,$$

由此递推不难看出 $\varepsilon_1,\cdots,\varepsilon_i$ 可由 $\alpha_1,\cdots,\alpha_i$ 线性表示，于是两个向量组等价. 即

$$L(\varepsilon_1,\cdots,\varepsilon_i)=L(\alpha_1,\cdots,\alpha_i) \quad (i=1,2,\cdots,s).$$

因为 $\alpha_1,\cdots,\alpha_i$ 线性无关，从上式可知，$\varepsilon_1,\cdots,\varepsilon_i$ 也线性无关，因而其中不会出现零向量，故$(\varepsilon_i,\varepsilon_i)\neq 0$. 在上面的公式中用它们作分母是有意义的.

(ii) 显然有

$$(\varepsilon_2,\varepsilon_1)=(\alpha_2,\varepsilon_1)-\frac{(\alpha_2,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}(\varepsilon_1,\varepsilon_1)=0.$$

利用数学归纳法不难证明：ε_{i+1}与每个 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_i$ 正交，从而 ε_1, $\varepsilon_2,\cdots,\varepsilon_s$ 两两正交.

如果在 V 中任取一组基

$$\alpha_1,\alpha_2,\cdots,\alpha_n.$$

利用施密特正交化方法求得

$$\varepsilon_1',\varepsilon_2',\cdots,\varepsilon_n',$$

这是 V 的一组基，两两正交. 只要把每个向量 ε_i' 单位化之后，即得 V 的一组标准正交基.

例 4 在欧氏空间 $\mathscr{R}^4$ 中取定一组基

$$\alpha_1=(1,1,0,0);\qquad \alpha_2(1,0,1,0);$$

$$\alpha_3=(-1,0,0,1);\quad \alpha_4=(1,-1,-1,1).$$

把它们正交化

$$\varepsilon_1'=\alpha_1=(1,1,0,0),$$

$$\varepsilon_2'=\alpha_2-\frac{(\alpha_2,\varepsilon_1')}{(\varepsilon_1',\varepsilon_1')}\varepsilon_1'=\left(\frac{1}{2},-\frac{1}{2},1,0\right),$$

$$\varepsilon_3'=\alpha_3-\frac{(\alpha_3,\varepsilon_1')}{(\varepsilon_1',\varepsilon_1')}\varepsilon_1'-\frac{(\alpha_3,\varepsilon_2')}{(\varepsilon_2',\varepsilon_2')}\varepsilon_2'=\left(-\frac{1}{3},\frac{1}{3},\frac{1}{3},1\right),$$

$$\varepsilon_4'=\alpha_4-\frac{(\alpha_4,\varepsilon_1')}{(\varepsilon_1',\varepsilon_1')}\varepsilon_1'-\frac{(\alpha_4,\varepsilon_2')}{(\varepsilon_2',\varepsilon_2')}\varepsilon_2'-\frac{(\alpha_4,\varepsilon_3')}{(\varepsilon_3',\varepsilon_3')}\varepsilon_3'$$
$$=(1,-1,-1,1).$$

再把每个向量单位化，得

$$\varepsilon_1=\frac{1}{|\varepsilon_1'|}\varepsilon_1'=\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0,0\right),$$

$$\varepsilon_2=\frac{1}{|\varepsilon_2'|}\varepsilon_2'=\left(\frac{1}{\sqrt{6}},-\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}},0\right),$$

$$\varepsilon_3=\frac{1}{|\varepsilon_3'|}\varepsilon_3'=\left(-\frac{1}{\sqrt{12}},\frac{1}{\sqrt{12}},\frac{1}{\sqrt{12}},\frac{3}{\sqrt{12}}\right),$$

$$\varepsilon_4=\frac{1}{|\varepsilon_4'|}\varepsilon_4'=\left(\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right).$$

这就得到 $\mathscr{R}^4$ 内的一组标准正交基. 如以它们为列向量(或行向量)排成一个四阶方阵

$$T=\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{12}} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & -\frac{1}{2} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & -\frac{1}{2} \\ 0 & 0 & \frac{3}{\sqrt{12}} & \frac{1}{2} \end{bmatrix},$$

那么,按命题 1.4,这是一个正交矩阵,$T'T=TT'=E$.

例 5 在 $\mathscr{R}[X]_4$ 内定义内积

$$(f,g)=\int_0^1 f(x)g(x)\mathrm{d}x$$

使之成一欧氏空间. 给定线性无关向量组

$$1,\ X,\ X^2,$$

把它正交化

$$\varepsilon'_1=1,$$

$$\varepsilon_2'=X-\frac{(X,\varepsilon_1')}{(\varepsilon_1',\varepsilon_1')}\varepsilon_1'=X-\frac{1}{2},$$

$$\varepsilon_3'=X^2-\frac{(X^2,\varepsilon_1')}{(\varepsilon_1',\varepsilon_1')}\varepsilon_1'-\frac{(X^2,\varepsilon_2')}{(\varepsilon_2',\varepsilon_2')}\varepsilon_2'$$

$$=X^2-X+\frac{1}{6}.$$

于是得到 $\mathscr{R}[X]_4$ 中与之等价的两两正交向量组

$$1,\ X-\frac{1}{2},\ X^2-X+\frac{1}{6}.$$

四、在标准正交基下内积的表达式

设 V 中取定一组标准正交基

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n,$$
$$(\varepsilon_i, \varepsilon_j) = \delta_{ij} \quad (i, j = 1, 2, \cdots, n).$$

命

$$\alpha = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n,$$
$$\beta = y_1\varepsilon_1 + y_2\varepsilon_2 + \cdots + y_n\varepsilon_n.$$

则有

$$(\alpha, \beta) = \sum_{i=1}^{n} \sum_{j=1}^{n} (\varepsilon_i, \varepsilon_j) x_i y_j = \sum_{i=1}^{n} \sum_{j=1}^{n} \delta_{ij} x_i y_j$$
$$= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

因此，在标准正交基下内积恰好等于两个向量的对应坐标相乘再相加. 这与三维几何空间中向量内积在直角坐标系下的计算公式相符.

*正　交　补

设 V 是一个 n 维欧氏空间，M 是它的一个子空间. 定义 V 的一个子集

$$M^{\perp} = \{\alpha \in V \mid \text{对一切 } \beta \in M \text{ 有} (\alpha, \beta) = 0\},$$

称 $M^{\perp}$ 为 M 的**正交补**. 显然，$M^{\perp}$ 关于 V 中向量的加法以及数乘运算是封闭的，故 $M^{\perp}$ 也是 V 的子空间.

命题 1.5　设 M 是 n 维欧氏空间 V 的一个子空间，则 V 可分解为 M 与 $M^{\perp}$ 的直和

$$V = M \oplus M^{\perp}.$$

证　设 $\alpha \in M \cap M^{\perp}$，由正交补的定义，有 $(\alpha, \alpha) = 0$. 所以 $\alpha = 0$. 这说明 $M \cap M^{\perp} = \{0\}$，即 $M + M^{\perp}$ 是直和. 取 M 的一组标准正交基 $\varepsilon_1, \cdots, \varepsilon_r$，扩充成 V 的一组标准正交基（根据施密特正交化方法，这总是可以办到的）

$$\varepsilon_1, \cdots, \varepsilon_r, \varepsilon_{r+1}, \cdots, \varepsilon_n.$$

易验证 $\varepsilon_{r+1}, \cdots, \varepsilon_n \in M^{\perp}$. 于是

$$V = L(\varepsilon_1, \cdots, \varepsilon_r) + L(\varepsilon_{r+1}, \cdots, \varepsilon_n) \subseteq M + M^{\perp} \subseteq V,$$

即 $V=M+M^{\perp}$. 所以 $V=M\oplus M^{\perp}$. ▌

推论 n 维欧氏空间 V 中任一两两正交单位向量组

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_s$$

都可扩充成 V 的一组标准正交基.

证 命 $M=L(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_s)$，则 $V=M\oplus M^{\perp}$. 再在 $M^{\perp}$ 内取一组基，按施密特正交化方法先把它正交化，然后再单位化，得 ε_{s+1}，$\cdots,\varepsilon_n$，则

$$\varepsilon_1,\cdots,\varepsilon_s,\varepsilon_{s+1},\cdots,\varepsilon_n$$

是 V 内一个两两正交的单位向量组，即为 V 的一组标准正交基. ▌

最后，我们简单介绍一下欧氏空间的同构概念.

定义 设 V_1,V_2 是两个欧氏空间. 如果存在 V_1 到 V_2 的一个映射 σ，满足如下条件：

(i) σ 是 V_1 到 V_2 的线性空间同构映射（即 $\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta)$，$\sigma(k\alpha)=k\sigma(\alpha)$，且 σ 是 V_1 到 V_2 的一一映射）；

(ii) σ 保持内积关系，即对任意 $\alpha,\beta\in V_1$，有

$$(\sigma(\alpha),\sigma(\beta))=(\alpha,\beta).$$

则称 σ 为欧氏空间 V_1 到欧氏空间 V_2 的**同构映射**. 此时称 V_1 与 V_2 **同构**.

同构的欧氏空间具有相同的代数性质和度量性质. 如果其中一个是有限维的，那么另一个也一定是有限维的，且两者维数相同.

现在设 V_1,V_2 是两个 n 维欧氏空间. 在 V_1 内取一组标准正交基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，在 V_2 内取一组标准正交基 $\varepsilon_1',\varepsilon_2',\cdots,\varepsilon_n'$. 定义 V_1 到 V_2 的映射 σ 如下：若

$$\alpha=a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n,$$

则令

$$\sigma(\alpha)=a_1\varepsilon_1'+a_2\varepsilon_2'+\cdots+a_n\varepsilon_n'.$$

容易验证，σ 是 V_1 到 V_2 的欧氏空间同构映射（具体验证留给读者作为练习）. 因此，对于有限维的欧氏空间来说，只要维数相同就彼

此同构. 特别地,任一 n 维欧氏空间都与欧氏空间 $\mathscr{R}^n$ 同构.

习 题 一

1. 设 A 是 n 阶正定矩阵. 在 $\mathscr{R}^n$ 中定义二元函数 (α,β) 如下: 若

$$\alpha=(x_1,x_2,\cdots,x_n),\quad \beta=(y_1,y_2,\cdots,y_n),$$

则令

$$(\alpha,\beta)=\alpha A\beta'.$$

证明:

(i) (α,β) 满足内积条件(i)～(iii),从而 $\mathscr{R}^n$ 关于这个内积也成一欧氏空间;

(ii) 写出这个欧氏空间的哥西-布尼雅可夫斯基不等式.

2. 在 $M_n(\mathscr{R})$ 中考虑全体 n 阶对称矩阵所成的子空间 V. 在 V 中定义二元函数如下:

$$(A,B)=\mathrm{Tr}(AB).$$

证明:这个函数满足内积条件,从而 V 关于它成一欧氏空间.

3. 在欧氏空间 $\mathscr{R}^4$ 中求向量 α,β 的夹角:

(1) $\alpha=(2,1,3,2),\beta=(1,2,-2,1)$;

(2) $\alpha=(1,2,2,3),\beta=(3,1,5,1)$;

(3) $\alpha=(1,1,1,2),\beta=(3,1,-1,0)$.

4. 证明:在欧氏空间中两向量 α,β 正交的充分必要条件是: 对任意实数 t,有

$$|\alpha+t\beta|\geqslant|\alpha|.$$

5. 在欧氏空间 V 内证明:

(1) $|\alpha+\beta|\leqslant|\alpha|+|\beta|$;

(2) 令 $d(\alpha,\beta)=|\alpha-\beta|$,则

$$d(\alpha,\gamma)\leqslant d(\alpha,\beta)+d(\beta,\gamma).$$

6. 在欧氏空间 $\mathscr{R}^4$ 中求一单位向量与

$$\alpha=(1,1,-1,1),\quad \beta=(1,-1,-1,1),\quad \gamma=(2,1,1,3)$$

正交.

7. 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是欧氏空间 V 的一组基,证明:

(1) 若 $(\beta,\alpha_i)=0(i=1,2,\cdots,n)$,则 $\beta=0$;

(2) 若 $(\beta_1,\alpha_i)=(\beta_2,\alpha_i)(i=1,2,\cdots,n)$,则 $\beta_1=\beta_2$.

8. 证明:欧氏空间 V 的任一子空间 M 关于 V 的内积也成一欧氏空间.

9. 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 是三维欧氏空间中一组标准正交基,证明:

$$\eta_1=\frac{1}{3}(2\varepsilon_1+2\varepsilon_2-\varepsilon_3),$$

$$\eta_2=\frac{1}{3}(2\varepsilon_1-\varepsilon_2+2\varepsilon_3),$$

$$\eta_3=\frac{1}{3}(\varepsilon_1-2\varepsilon_2-2\varepsilon_3)$$

也是一组标准正交基.

10. 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4,\varepsilon_5$ 是五维欧氏空间 V 的一组标准正交基,而

$$\alpha_1=\varepsilon_1+\varepsilon_5,\quad \alpha_2=\varepsilon_1-\varepsilon_2+\varepsilon_4,\quad \alpha_3=2\varepsilon_1+\varepsilon_2+\varepsilon_3,$$

求 $L(\alpha_1,\alpha_2,\alpha_3)$的一组标准正交基.

11. 求齐次线性方程组

$$\begin{cases}2x_1+x_2-x_3+x_4-3x_5=0,\\ x_1+x_2-x_3\qquad\quad+\ x_5=0\end{cases}$$

的解空间(作为欧氏空间 $\mathscr{R}^5$ 的子空间)的一组标准正交基.

12. 考虑欧氏空间 $C[-\pi,\pi]$(参看本节例 2),在其中给定向量组

$$1,\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x,\ \cdots,\ \cos nx,\ \sin nx.$$

证明:

(1) 这个向量组的向量两两正交;

(2) 求这个向量组生成的子空间 M 的一组标准正交基.

13. 在 $\mathscr{R}[X]_4$ 中定义内积

$$(f,g)=\int_{-1}^{1} f(x)g(x)\mathrm{d}x,$$

试求出它的一组标准正交基.

14. 在欧氏空间 $\mathscr{R}^4$ 内给定向量组 $\alpha_1,\alpha_2,\alpha_3$，利用施密特正交化方法先把它正交化，再单位化.

(1) $\alpha_1=(1,2,-1,0)$； $\alpha_2=(1,-1,1,1)$；

$\alpha_3=(-1,2,1,1)$.

(2) $\alpha_1=(2,1,0,1)$； $\alpha_2=(0,1,2,2)$；

$\alpha_3=(-2,1,1,2)$.

15. 设 V 是 n 维欧氏空间，α 是 V 中一固定向量.

(1) 证明：

$$M=\{\beta\in V\mid(\beta,\alpha)=0\}$$

是 V 的一个子空间；

(2) 证明当 $\alpha\neq0$ 时 $\dim M=n-1$.

16. 设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是欧氏空间 V 内一个向量组，令

$$D=\begin{vmatrix}(\alpha_1,\alpha_1) & (\alpha_1,\alpha_2) & \cdots & (\alpha_1,\alpha_s)\\ (\alpha_2,\alpha_1) & (\alpha_2,\alpha_2) & \cdots & (\alpha_2,\alpha_s)\\ \vdots & \vdots & & \vdots\\ (\alpha_s,\alpha_1) & (\alpha_s,\alpha_2) & \cdots & (\alpha_s,\alpha_s)\end{vmatrix}.$$

证明：$\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关的充分必要条件是 $D\neq0$.

17. 给定两个四维向量

$$\alpha_1=\left(\frac{1}{3},-\frac{2}{3},0,\frac{2}{3}\right),$$

$$\alpha_2=\left(-\frac{2}{\sqrt6},0,\frac{1}{\sqrt6},\frac{1}{\sqrt6}\right),$$

求作一个四阶正交矩阵 T，以 α_1,α_2 作为它的前两个列向量.

18. 证明：实上三角矩阵为正交矩阵时，必为对角矩阵，且对角线上的元素为 ±1.

19. 设 A 是一个 n 阶实方阵，$|A|\neq0$. 证明 A 可分解为一个

正交矩阵 Q 和一个上三角矩阵

$$T=\begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ & t_{22} & \cdots & t_{2n} \\ & & \ddots & \vdots \\ & 0 & & t_{nn} \end{pmatrix} \quad (t_{ii}>0, i=1,2,\cdots,n)$$

的乘积：$A=QT$. 并证明这种分解是唯一的.

20. 设 A 是 n 阶正定矩阵，证明存在一个上三角矩阵 T，使 $A=T'T$.

21. 设 $f(\alpha)$是 n 维欧氏空间 V 内的一个线性函数，证明在 V 内存在一个固定向量 β，使对一切 $\alpha\in V$，有

$$f(\alpha)=(\alpha,\beta).$$

§2 正 交 变 换

在这一节里，我们研究欧氏空间中的一类重要的线性变换.

定义 设 $\boldsymbol{A}$ 是欧氏空间 V 中的一个线性变换，且对任意 $\alpha,\beta\in V$，有

$$(\boldsymbol{A}\alpha,\boldsymbol{A}\beta)=(\alpha,\beta),$$

则称 $\boldsymbol{A}$ 是一个**正交变换**.

由于正交变换保持向量的内积不变，因而它保持向量的长度和夹角不变，即

$$|\boldsymbol{A}\alpha|=|\alpha|,$$

$$<\boldsymbol{A}\alpha,\boldsymbol{A}\beta>=<\alpha,\beta>,\quad \alpha\neq 0,\beta\neq 0.$$

特别地，若 $\alpha\perp\beta$，则 $\boldsymbol{A}\alpha\perp\boldsymbol{A}\beta$.

命题 2.1 设 $\boldsymbol{A}$ 是 n 维欧氏空间 V 内的一个线性变换，则下列命题互相等价：

(i) $\boldsymbol{A}$ 是正交变换；

(ii) 对任意 $\alpha\in V$，有 $|\boldsymbol{A}\alpha|=|\alpha|$；

(iii) 若 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组标准正交基，则 $\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,$

$\boldsymbol{A}\varepsilon_n$ 也是 V 的一组标准正交基；

(iv) $\boldsymbol{A}$ 在 V 的某一组标准正交基下的矩阵是正交矩阵.

证 采用轮转的证法.

(i)⇒(ii)：显然.

(ii)⇒(iii)：按假设，有

$$|\boldsymbol{A}\varepsilon_i|=|\varepsilon_i|=1 \quad (i=1,2,\cdots,n).$$

下面只要证$(\boldsymbol{A}\varepsilon_i,\boldsymbol{A}\varepsilon_j)=0(i\neq j)$即可. 因

$$|\boldsymbol{A}(\varepsilon_i+\varepsilon_j)|=|\varepsilon_i+\varepsilon_j|,$$

所以

$$(\boldsymbol{A}(\varepsilon_i+\varepsilon_j),\boldsymbol{A}(\varepsilon_i+\varepsilon_j))=(\varepsilon_i+\varepsilon_j,\varepsilon_i+\varepsilon_j).$$

展开，得

$$(\boldsymbol{A}\varepsilon_i,\boldsymbol{A}\varepsilon_i)+2(\boldsymbol{A}\varepsilon_i,\boldsymbol{A}\varepsilon_j)+(\boldsymbol{A}\varepsilon_j,\boldsymbol{A}\varepsilon_j)$$
$$=(\varepsilon_i,\varepsilon_i)+2(\varepsilon_i,\varepsilon_j)+(\varepsilon_j,\varepsilon_j).$$

故有

$$(\boldsymbol{A}\varepsilon_i,\boldsymbol{A}\varepsilon_j)=(\varepsilon_i,\varepsilon_j)=0 \quad (i\neq j).$$

(iii)⇒(iv)：若 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组标准正交基，则 $\boldsymbol{A}\varepsilon_1$，$\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n$ 也是一组标准正交基. 而

$$(\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A.$$

现在 A 既是线性变换 $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵，又是这两组标准正交基之间的过渡矩阵，由命题 1.3 知，A 是一个正交矩阵.

(iv)⇒(i)：设 $\boldsymbol{A}$ 在标准正交基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵 A 是正交矩阵，由命题 1.3 知，$\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n$ 是 V 的一组标准正交基. 对任意 $\alpha,\beta\in V$，命

$$\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n,$$
$$\beta=y_1\varepsilon_1+y_2\varepsilon_2+\cdots+y_n\varepsilon_n.$$

以 $\boldsymbol{A}$ 作用于上式两端，得

$$\boldsymbol{A}\alpha=x_1\boldsymbol{A}\varepsilon_1+x_2\boldsymbol{A}\varepsilon_2+\cdots+x_n\boldsymbol{A}\varepsilon_n,$$
$$\boldsymbol{A}\beta=y_1\boldsymbol{A}\varepsilon_1+y_2\boldsymbol{A}\varepsilon_2+\cdots+y_n\boldsymbol{A}\varepsilon_n.$$

所以

$$(\boldsymbol{A}\alpha,\boldsymbol{A}\beta)=x_1y_1+x_2y_2+\cdots+x_ny_n=(\alpha,\beta).$$

即 $\boldsymbol{A}$ 是正交变换. ▎

命题 2.1 给出了正交变换的四种互相等价的说法，其中每一种都可以用作正交变换的定义.

设 V 是一个 n 维欧氏空间，命 $O(n)$ 表示 V 中全体正交变换所成的集合. 我们有

命题 2.2 $O(n)$ 具有如下性质：

(i) $\boldsymbol{E}\in O(n)$；

(ii) 若 $\boldsymbol{A},\boldsymbol{B}\in O(n)$，则 $\boldsymbol{AB}\in O(n)$；

(iii) 若 $\boldsymbol{A}\in O(n)$，则 $\boldsymbol{A}$ 可逆，且 $\boldsymbol{A}^{-1}\in O(n)$.

证 (i)是显然的. 只证(ii)与(iii).

(ii)：对任意 $\alpha,\beta\in V$，有

$$(\boldsymbol{AB}\alpha,\boldsymbol{AB}\beta)=(\boldsymbol{A}(\boldsymbol{B}\alpha),\boldsymbol{A}(\boldsymbol{B}\beta))=(\boldsymbol{B}\alpha,\boldsymbol{B}\beta)=(\alpha,\beta).$$

故 $\boldsymbol{AB}$ 是正交变换.

(iii)：正交变换 $\boldsymbol{A}$ 在任一组标准正交基下的矩阵都是正交矩阵，而正交矩阵都可逆，故 $\boldsymbol{A}$ 也可逆. 对任意 $\alpha,\beta\in V$，有

$$\begin{aligned}(\alpha,\beta)&=(\boldsymbol{E}\alpha,\boldsymbol{E}\beta)=(\boldsymbol{A}(\boldsymbol{A}^{-1}\alpha),\boldsymbol{A}(\boldsymbol{A}^{-1}\beta))\\&=(\boldsymbol{A}^{-1}\alpha,\boldsymbol{A}^{-1}\beta).\end{aligned}$$

即 $\boldsymbol{A}^{-1}$是一正交变换. ▎

一个非空的集合 G，如果其元素之间存在着乘法运算，且这种乘法运算满足结合律，又具有命题 2.2 中所列举出的性质，那么，在代数学中把这样的 G 称为一个**群**(这里不给出其严格的定义). 因此，$O(n)$(关于线性变换的乘法)是一个群，称为 n 阶**正交群**，它在理论物理等领域中具有重要的意义.

现设正交变换 $\boldsymbol{A}$ 在 V 的一组标准正交基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵为 A，则 A 是正交矩阵，$AA'=E$，故

$$|AA'|=|A|\cdot|A'|=|A|^2=|E|=1.$$

于是 $|A|=\pm1$. $\boldsymbol{A}$ 在不同基下的矩阵是相似的，而相似矩阵具有

相同的行列式，所以，如果 $\boldsymbol{A}$ 在某一组标准正交基下的矩阵的行列式为 $+1$，则它在所有标准正交基下的矩阵的行列式都是 $+1$，此时 $\boldsymbol{A}$ 称为**第一类正交变换**（或称为 V 内的一个**旋转**）. 如 $|A|=-1$，则 $\boldsymbol{A}$ 称为**第二类正交变换**.

习 题 二

1. 设 $\boldsymbol{A}$ 是欧氏空间 V 内的线性变换，不用命题 2.1，直接证明：若 $\boldsymbol{A}$ 保持 V 中向量的长度不变，则它保持向量的内积不变（从而也保持两个非零向量间的夹角不变）. 这样，$\boldsymbol{A}$ 就是 V 内的一个正交变换，

2. 设 $\boldsymbol{A}$ 是欧氏空间 V 内的一个正交变换，证明：若 $\boldsymbol{A}^2=\boldsymbol{E}$，则 $(\boldsymbol{A}\alpha,\beta)=(\alpha,\boldsymbol{A}\beta)$.

3. 证明正交变换的特征值只能是 ± 1.

4. 设 η 是 n 维欧氏空间 V 内的一个单位向量，定义 V 内一个线性变换如下：

$$\boldsymbol{A}\alpha=\alpha-2(\eta,\alpha)\eta \quad (\alpha\in V),$$

称这样的线性变换 $\boldsymbol{A}$ 为一个**镜面反射**. 证明：

(1) $\boldsymbol{A}$ 是正交变换；

(2) $\boldsymbol{A}$ 是第二类的；

(3) $\boldsymbol{A}^2=\boldsymbol{E}$；

(4) 设 $\boldsymbol{B}$ 是 V 内一个第二类正交变换，则必有

$$\boldsymbol{B}=\boldsymbol{A}\cdot\boldsymbol{B}_1,$$

其中 $\boldsymbol{B}_1$ 是 V 内的一个第一类正交变换.

5. 设 V 是一个 n 维欧氏空间，V 中一个正交变换 $\boldsymbol{A}$ 有特征值 $\lambda_0=1$，且 $\dim V_{\lambda_0}=n-1$，证明 $\boldsymbol{A}$ 是一个镜面反射.

6. 设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 和 $\beta_1,\beta_2,\cdots,\beta_s$ 是 n 维欧氏空间 V 中两个向量组，证明存在一个正交变换 $\boldsymbol{A}$，使

$$\boldsymbol{A}\alpha_i=\beta_i \quad (i=1,2,\cdots,s)$$

的充分必要条件是

$$(\alpha_i,\alpha_j)=(\beta_i,\beta_j)\quad(i,j=1,2,\cdots,s).$$

7. 证明:奇数维欧氏空间中的旋转必有一个特征值 $\lambda_0=1$.

8. 证明:第二类正交变换必有一个特征值 $\lambda_0=-1$.

9. 设 α,β 是欧氏空间中两个不同的单位向量,证明存在一个镜面反射 $\boldsymbol{A}$,使 $\boldsymbol{A}\alpha=\beta$.

*10. 证明:n 维欧氏空间中任一正交变换都可以表成一系列镜面反射的乘积.

11. 设 $\boldsymbol{A}$ 是 n 维欧氏空间 V 内的一个正交变换,M 是 $\boldsymbol{A}$ 的一个不变子空间,证明 M 的正交补 $M^{\perp}$ 也是 $\boldsymbol{A}$ 的不变子空间.

*12. 设 $\boldsymbol{A}$ 是欧氏空间 V 内的一个变换,对任意 $\alpha,\beta\in V$,有 $(\boldsymbol{A}\alpha,\boldsymbol{A}\beta)=(\alpha,\beta)$,证明 $\boldsymbol{A}$ 是一个正交变换.

§3 对称变换

在这一节里,我们研究欧氏空间中另一类重要的线性变换,

定义 设 $\boldsymbol{A}$ 是 n 维欧氏空间 V 内的一个线性变换,如果对任意 $\alpha,\beta\in V$,有

$$(\boldsymbol{A}\alpha,\beta)=(\alpha,\boldsymbol{A}\beta),$$

则称 $\boldsymbol{A}$ 是一个**对称变换**.

下面的命题说明了把这种线性变换称为对称变换的理由.

命题 3.1 设 $\boldsymbol{A}$ 是 n 维欧氏空间 V 内的一个线性变换. 则有:

(i) 如果 $\boldsymbol{A}$ 是一个对称变换,那么,它在 V 的任一标准正交基下的矩阵都是实对称矩阵;

(ii) 如果 $\boldsymbol{A}$ 在 V 的某一组标准正交基下的矩阵是实对称矩阵,则 $\boldsymbol{A}$ 是对称变换.

证 分别证明这两条结论.

(i) 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组标准正交基. 又设

$$(\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A,$$

其中 $A=(a_{ij})$. 于是

$$\boldsymbol{A}\varepsilon_i = a_{1i}\varepsilon_1 + a_{2i}\varepsilon_2 + \cdots + a_{ni}\varepsilon_n.$$

$$(\boldsymbol{A}\varepsilon_i, \varepsilon_j) = a_{ji}(\varepsilon_j, \varepsilon_j) = a_{ji},$$

$$(\varepsilon_i, \boldsymbol{A}\varepsilon_j) = a_{ij}(\varepsilon_i, \varepsilon_i) = a_{ij}.$$

即有

$$a_{ji} = (\boldsymbol{A}\varepsilon_i, \varepsilon_j) = (\varepsilon_i, \boldsymbol{A}\varepsilon_j) = a_{ij} \quad (i, j = 1, 2, \cdots, n),$$

故 A 为实对称矩阵.

(ii)设 $\boldsymbol{A}$ 在标准正交基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵 $A = (a_{ij})$是实对称矩阵. 与(i)的证明相同,有

$$(\boldsymbol{A}\varepsilon_i, \varepsilon_j) = a_{ji}, \quad (\varepsilon_i, \boldsymbol{A}\varepsilon_j) = a_{ij}.$$

故

$$(\boldsymbol{A}\varepsilon_i, \varepsilon_j) = (\varepsilon_i, \boldsymbol{A}\varepsilon_j) \quad (i, j = 1, 2, \cdots, n). \tag{1}$$

对任意 $\alpha, \beta \in V$,设

$$\begin{cases} \alpha = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n, \\ \beta = y_1\varepsilon_1 + y_2\varepsilon_2 + \cdots + y_n\varepsilon_n. \end{cases}$$

则

$$(\boldsymbol{A}\alpha, \beta) = \left(\sum_{k=1}^{n} x_k \boldsymbol{A}\varepsilon_k, \sum_{l=1}^{n} y_l \varepsilon_l \right) = \sum_{k=1}^{n} \sum_{l=1}^{n} (\boldsymbol{A}\varepsilon_k, \varepsilon_l) x_k y_l.$$

$$(\alpha, \boldsymbol{A}\beta) = \left(\sum_{k=1}^{n} x_k \varepsilon_k, \sum_{l=1}^{n} y_l \boldsymbol{A}\varepsilon_l \right) = \sum_{k=1}^{n} \sum_{l=1}^{n} (\varepsilon_k, \boldsymbol{A}\varepsilon_l) x_k y_l.$$

根据(1)式,有$(\boldsymbol{A}\alpha, \beta) = (\alpha, \boldsymbol{A}\beta)$,即 $\boldsymbol{A}$ 为对称变换. ▌

命题 3.1 说明,当取定 V 的一组标准正交基之后,V 中的全体对称变换所成的集合和 n 阶实对称矩阵所成的集合之间就可以建立起一一对应的关系. 因而,在研究对称变换时可以利用实对称矩阵的性质;反之,研究实对称矩阵时也可以利用对称变换的结果,这两者相辅相成.

对称变换的基本定理

我们证明：对 n 维欧氏空间 V 内的一个对称变换 $\boldsymbol{A}$，我们一定可以找出一组标准正交基，使 $\boldsymbol{A}$ 在这组基下的矩阵成对角形．这就是对称变换的基本定理．

命题 3.2 实对称矩阵 A 的特征值都是实数．

证 A 的特征多项式 $f(\lambda)=|\lambda E-A|$ 在复数域内的任一根设为 λ_0．我们证明 λ_0 必为实数．

设 X 为复的 n 维列向量，$X\neq 0$，满足

$$AX=\lambda_0 X. \tag{2}$$

两边取复共轭，再取转置．因为 $\overline{A}=A$（$\overline{A}$ 是把 A 的每个元素取复数共轭所得的矩阵，$\overline{X}$ 亦然），$A'=A$，有

$$\overline{X}'A=\overline{\lambda}_0\overline{X}'.$$

两边以 X 右乘，得

$$\overline{X}'AX=\overline{\lambda}_0\overline{X}'X. \tag{3}$$

再以 $\overline{X}'$ 左乘等式(2)，得

$$\overline{X}'AX=\lambda_0\overline{X}'X. \tag{4}$$

比较(3)与(4)，得

$$\overline{\lambda}_0\overline{X}'X=\lambda_0\overline{X}'X. \tag{5}$$

因为 $\overline{X}'X$ 是 X 分量的模的平方和，即

$$\overline{X}'X=|x_1|^2+|x_2|^2+\cdots+|x_n|^2,$$

而 $X\neq 0$，故 $\overline{X}'X\neq 0$．由(5)式即得 $\lambda_0=\overline{\lambda}_0$，于是，$\lambda_0$ 为实数．▌

推论 欧氏空间 V 内任一对称变换 $\boldsymbol{A}$ 至少有一个特征值．

证 $\boldsymbol{A}$ 在某一组标准正交基下的矩阵 A 为实对称矩阵，由命题 3.2 知，A 的特征值全是实数，因而全是 $\boldsymbol{A}$ 的特征值．故 $\boldsymbol{A}$ 至少有一个特征值．▌

命题 3.3 设 $\boldsymbol{A}$ 是欧氏空间 V 内的一个对称变换，则 $\boldsymbol{A}$ 的对应于不同特征值 λ_1,λ_2 的特征向量 ξ_1,ξ_2 互相正交．

证 根据命题的条件，$\lambda_1\neq\lambda_2$，且

$$A\xi_1=\lambda_1\xi_1,\quad A\xi_2=\lambda_2\xi_2.$$

于是 $$\lambda_1(\xi_1,\xi_2)=(A\xi_1,\xi_2)=(\xi_1,A\xi_2)=\lambda_2(\xi_1,\xi_2).$$

移项,得

$$(\lambda_1-\lambda_2)(\xi_1,\xi_2)=0.$$

因 $\lambda_1-\lambda_2\neq 0$,故 $(\xi_1,\xi_2)=0$. ▌

命题 3.4 设 A 是 n 维欧氏空间 V 内的一个对称变换,λ_0 是 A 的一个特征值,ξ 是对应于 λ_0 的特征向量. 命

$$M=\{\alpha\in V\mid(\alpha,\xi)=0\},$$

则 M 是 A 的不变子空间,且 $\dim M=n-1$.

证 对任一 $\alpha\in M$,有

$$(A\alpha,\xi)=(\alpha,A\xi)=\lambda_0(\alpha,\xi)=0.$$

即 $A\alpha\in M$. 故 M 为 A 的不变子空间. 把 ξ 扩充为 V 的一组基

$$\xi_1=\xi,\ \xi_2,\ \cdots,\ \xi_n.$$

利用施密特正交化方法先把它正交化,得到 V 的一组两两正交的基

$$\varepsilon_1,\ \varepsilon_2,\ \cdots,\ \varepsilon_n.$$

其中 $\varepsilon_1=\xi_1$. 显然,$\varepsilon_2,\cdots,\varepsilon_n\in M$,所以 $L(\varepsilon_2,\cdots,\varepsilon_n)\subseteq M$. 反之,对任一 $\alpha\in M$,设 $\alpha=a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n$,则 $0=(\alpha,\varepsilon_1)=a_1(\varepsilon_1,\varepsilon_1)$,由此得 $a_1=0$,即 $\alpha\in L(\varepsilon_2,\cdots,\varepsilon_n)$,于是 $M=L(\varepsilon_2,\cdots,\varepsilon_n)$. 故 $\dim M=n-1$. ▌

下面我们证明本章的主要定理.

定理 1 设 A 是 n 维欧氏空间 V 内的一个对称变换,则在 V 内存在一组标准正交基,使 A 在此组基下的矩阵成对角形.

证 对 V 的维数 n 作数学归纳法.

$n=1$ 时命题是显然的. 现设命题对 $n-1$ 维的欧氏空间成立,证明它对 n 维欧氏空间也成立.

从命题 3.2 的推论知:A 在 V 内必有一特征值 λ_1. 设对应于 λ_1 的单位特征向量为 η_1,即

$$A\eta_1=\lambda_1\eta_1,\quad(\eta_1,\eta_1)=1.$$

命

$$M=\{\alpha\in V\mid(\alpha,\eta_1)=0\}.$$

由命题 3.4 知，M 是 $\boldsymbol{A}$ 的 $n-1$ 维不变子空间. M 关于 V 的内积也是一个欧氏空间，$\boldsymbol{A}$ 限制在 M 内也是一个对称变换. 根据归纳假设，在 M 内存在一组标准正交基 $\eta_2,\eta_3,\cdots,\eta_n$，使

$$\boldsymbol{A}\eta_i=\lambda_i\eta_i\quad(i=2,3,\cdots,n).$$

现在向量组 $\eta_1,\eta_2,\cdots,\eta_n$ 满足

$$(\eta_i,\eta_j)=\delta_{ij}\quad(i,j=1,2,\cdots,n),$$

$$\boldsymbol{A}\eta_i=\lambda_i\eta_i\quad(i=1,2,\cdots,n),$$

故它是 V 的一组标准正交基，且 $\boldsymbol{A}$ 在这组基下的矩阵成对角形. ▌

推论 设 A 是一个 n 阶实对称矩阵，则存在 n 阶正交矩阵 T，使

$$T^{-1}AT=T'AT=D$$

为对角矩阵.

证 把 A 看作 n 维欧氏空间 V 内一个对称变换 $\boldsymbol{A}$ 在标准正交基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵. 从定理 1 知，在 V 内存在标准正交基 $\eta_1,\eta_2,\cdots,\eta_n$，使 $\boldsymbol{A}$ 在这组基下的矩阵成对角形 D. 令

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T;$$

则 T 为正交矩阵，且 $T^{-1}AT=T'AT=D$. ▌

定理 2 给定 n 个未知量 $x_1,x_2,\cdots,x_n$ 的实二次型

$$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\quad(a_{ij}=a_{ji}),$$

则存在一个 n 阶正交矩阵 T，使在线性变数替换 $X=TZ$ 下二次型化为标准形

$$\lambda_1z_1^2+\lambda_2z_2^2+\cdots+\lambda_nz_n^2.$$

且 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 除了可能差一个排列次序外，是被 f 唯一确定的.

证 二次型 f 的矩阵 $A=(a_{ij})$ 是一个实对称矩阵，根据定理 1 的推论，存在正交矩阵 T，使

$$T'AT=D=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}.$$

而

$$\begin{aligned}f&=X'AX\xlongequal{X=TZ}(TZ)'A(TZ)\\&=Z'(T'AT)Z=Z'DZ\\&=\lambda_1z_1^2+\lambda_2z_2^2+\cdots+\lambda_nz_n^2,\end{aligned}$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是矩阵 A 的全部特征值，除了排列次序可任意外，是由 f 唯一确定的. ▌

为着讨论正交矩阵 T 的实际计算方法，我们先证如下的命题.

命题 3.5 设 $\boldsymbol{A}$ 是 n 维欧氏空间 V 内的一个对称变换，$\lambda_1,\lambda_2,\cdots,\lambda_s$ 是 $\boldsymbol{A}$ 的全部互不相同的特征值，则从每个特征子空间 $V_{\lambda_i}(i=1,2,\cdots,s)$ 中各取一组标准正交基，即可并成 V 的一组标准正交基 $\eta_1,\eta_2,\cdots,\eta_n$，$\boldsymbol{A}$ 在这组基下的矩阵为对角形.

证 首先假定这样并成一个向量组 $\eta_1,\eta_2,\cdots,\eta_m$. 由命题 3.3 知，这是一个两两正交的单位向量组. 又根据命题 1.3，它们线性无关. 其次，从定理 1 知，V 内存在一组由 $\boldsymbol{A}$ 的特征向量所组成的标准正交基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$. 因为每个 ε_i 必属于某个 V_{λ_j}，故 ε_i 可被 $\eta_1,\eta_2,\cdots,\eta_m$ 线性表示. 利用第四章命题 2.1 的(ii)知，$m=n$，且 $\eta_1,\eta_2,\cdots,\eta_n$ 构成 V 的一组标准正交基. 因为诸 η_i 均为 $\boldsymbol{A}$ 的特征向量，故 $\boldsymbol{A}$ 在这组基下的矩阵成对角形. ▌

根据这一命题，只要把 $\boldsymbol{A}$ 的全部不同特征值找出来，对每个 λ_i 找出 V_{λ_i} 的一组基(这相当于求一个齐次线性方程组的一个基础解系)，再利用施密特正交化方法把它们分别正交化再单位化，然后并起来，就得到定理 1 中所说的标准正交基.

推论 设 $\boldsymbol{A}$ 是 n 维欧氏空间 V 内的一个对称变换，λ_0 是 $\boldsymbol{A}$ 的一个 k 重特征值，则 $\dim V_{\lambda_0}=k$.

证 设 $\boldsymbol{A}$ 的全部不同特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_s$，其重数分别为

$n_1, n_2, \cdots, n_s$，则 $n_1+n_2+\cdots+n_s=n$. 从命题 3.5 知，$\dim V_{\lambda_i} \leqslant n_i (i=1,2,\cdots,s)$. 但

$$\dim V_{\lambda_1}+\dim V_{\lambda_2}+\cdots+\dim V_{\lambda_s}=n,$$

故

$$\dim V_{\lambda_i}=n_i \quad (i=1,2,\cdots,s). \blacksquare$$

用正交矩阵化实对称矩阵成对角形

给定 n 阶实对称矩阵 A，我们研究如何找出正交矩阵 T，使 $T'AT$ 成对角形. 我们的办法是：把 A 看作某个 n 维欧氏空间中的一个对称变换在一组标准正交基下的矩阵，然后再使用前面从理论上获得的一般性结果来探讨 T 的计算法.

为此，考虑欧氏空间 $\mathscr{R}^n$，将其向量表成 $n\times 1$ 矩阵形式. 定义其线性变换如下：

$$\boldsymbol{A}X=AX \quad (X\in \mathscr{R}^n).$$

$\mathscr{R}^n$ 内有一组标准正交基

$$\varepsilon_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},\quad \varepsilon_2=\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\quad \cdots,\quad \varepsilon_n=\begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}.$$

显然，$\boldsymbol{A}$ 在这组基下的矩阵即为 A，故 $\boldsymbol{A}$ 是一个对称变换. 根据定理 1 知，在 $\mathscr{R}^n$ 内存在一组标准正交基

$$\eta_1=\begin{pmatrix}t_{11}\\t_{21}\\\vdots\\t_{n1}\end{pmatrix},\quad \eta_2=\begin{pmatrix}t_{12}\\t_{22}\\\vdots\\t_{n2}\end{pmatrix},\quad \cdots,\quad \eta_n=\begin{pmatrix}t_{1n}\\t_{2n}\\\vdots\\t_{nn}\end{pmatrix}$$

使

$$\boldsymbol{A}\eta_i=A\eta_i=A\begin{pmatrix}t_{1i}\\t_{2i}\\\vdots\\t_{ni}\end{pmatrix}=\lambda_i\begin{pmatrix}t_{1i}\\t_{2i}\\\vdots\\t_{ni}\end{pmatrix}. \tag{6}$$

以 $\eta_1,\eta_2,\cdots,\eta_n$ 作列向量排成一个 n 阶方阵

$$T=\begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{pmatrix}.$$

从命题 1.4 可知,T 是一个正交矩阵. 把(6)式合并起来写成如下矩阵乘积形式

$$AT=T\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}, \quad D=\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

于是有 $T'AT=D$ 为对角矩阵(或者这样看,T 恰为 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 到 $\eta_1,\eta_2,\cdots,\eta_n$ 的过渡矩阵,$\boldsymbol{A}$ 在这两组基下的矩阵分别为 A 与 D,从而 $T^{-1}AT=T'AT=D$).

根据以上的分析,再根据命题 3.5,我们得到矩阵 T 的计算方法如下:

1. 计算特征多项式 $f(\lambda)=|\lambda E-A|$,求出它的全部互不相同的特征根(全是实数)$\lambda_1,\lambda_2,\cdots,\lambda_s$.

2. 对每个 λ_i 求出 V_{λ_i} 的一组基. 这就是求齐次线性方程组

$$(\lambda_i E-A)X=0$$

的一个基础解系

$$X_{i1},\ X_{i2},\ \cdots,\ X_{in_i}\quad (i=1,2,\cdots,s).$$

3. 利用施密特正交化方法,在 $\mathscr{R}^n$ 内把上面得到的每一个基础解系正交化后再单位化,得 V_{λ_i} 的一组标准正交基

$$\eta_{i1},\ \eta_{i2},\ \cdots,\ \eta_{in_i}\ (i=1,2,\cdots,s).$$

4. 把所得 s 个正交单位向量组作为列向量组排成一个 n 阶方阵 T,T 即为所求的正交矩阵. 此时要注意 D 的主对角线上 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 的排列次序要与 η_{ij} 的排列次序相对应.

例 1 给定实对称矩阵

$$A=\begin{pmatrix}0 & 1 & 1 & -1\\ 1 & 0 & -1 & 1\\ 1 & -1 & 0 & 1\\ -1 & 1 & 1 & 0\end{pmatrix},$$

求正交矩阵 T，使 $T'AT$ 成对角形.

解 (i) 求 A 的全部特征值.

$$|\lambda E-A|=\begin{vmatrix}\lambda & -1 & -1 & 1\\ -1 & \lambda & 1 & -1\\ -1 & 1 & \lambda & -1\\ 1 & -1 & -1 & \lambda\end{vmatrix}=(\lambda-1)^3(\lambda+3).$$

故 A 的互不相同的特征值为 $\lambda_1=1,\lambda_2=-3$.

(ii) 求每个特征值对应的线性无关特征向量.

$\lambda_1=1$：

$$\lambda_1 E-A=\begin{pmatrix}1 & -1 & -1 & 1\\ -1 & 1 & 1 & -1\\ -1 & 1 & 1 & -1\\ 1 & -1 & -1 & 1\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1 & -1 & -1 & 1\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}$$

$$\Longleftrightarrow x_1-x_2-x_3+x_4=0.$$

移项，得

$$x_1=x_2+x_3-x_4.$$

基础解系为(此时向量改写为行的形式，下面作正交化时较为方便)

$$X_{11}=(1,1,0,0),\ X_{12}=(1,0,1,0),$$

$$X_{13}=(-1,0,0,1).$$

$\lambda_2=-3$：

$$\lambda_2 E - A = \begin{pmatrix} -3 & -1 & -1 & 1 \\ -1 & -3 & 1 & -1 \\ -1 & 1 & -3 & -1 \\ 1 & -1 & -1 & -3 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & -1 & -1 & -3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\Longleftrightarrow \begin{cases} x_1 - x_2 - x_3 - 3x_4 = 0, \\ \quad x_2 \quad + x_4 = 0, \\ \quad\quad x_3 + x_4 = 0. \end{cases}$$

它的基础解系是 $\quad X_{21} = (1, -1, -1, 1).$

(iii) 把 X_{11}, X_{12}, X_{13} 正交化

$$\alpha_1 = X_{11} = (1, 1, 0, 0),$$

$$\alpha_1 = X_{12} - \frac{(X_{12}, \alpha_1)}{(\alpha_1, \alpha_1)}\alpha_1 = \left(\frac{1}{2}, -\frac{1}{2}, 1, 0\right),$$

$$\alpha_3 = X_{13} - \frac{(X_{13}, \alpha_1)}{(\alpha_1, \alpha_1)}\alpha_1 - \frac{(X_{13}, \alpha_2)}{(\alpha_2, \alpha_2)}\alpha_2$$

$$= \left(-\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 1\right).$$

再把 $\alpha_1, \alpha_2, \alpha_3$ 和 X_{21} 分别单位化：

$$\eta_1 = \frac{1}{|\alpha_1|}\alpha_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right),$$

$$\eta_2 = \frac{1}{|\alpha_2|}\alpha_2 = \left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0\right),$$

$$\eta_3 = \frac{1}{|\alpha_3|}\alpha_3 = \left(-\frac{1}{\sqrt{12}}, \frac{1}{\sqrt{12}}, \frac{1}{\sqrt{12}}, \frac{3}{\sqrt{12}}\right),$$

$$\eta_4 = \frac{1}{|X_{21}|}X_{21} = \left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right).$$

(iv) 以 $\eta_1, \eta_2, \eta_3, \eta_4$ 为列向量组排成矩阵

$$T=\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{12}} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & -\frac{1}{2} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & -\frac{1}{2} \\ 0 & 0 & \frac{3}{\sqrt{12}} & \frac{1}{2} \end{bmatrix},$$

则有

$$T'AT=D=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}.$$

下面，我们讨论如何将上面所得的结果应用于实二次型

$$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j \quad (a_{ij}=a_{ji}).$$

f 的矩阵为 $A=(a_{ij})$. 因 $T'AT=D$ 为对角矩阵，故二次型 f 在可逆线性变数替换 $X=TZ$ 下化为标准形

$$\lambda_1z_1^2+\lambda_2z_2^2+\cdots+\lambda_nz_n^2.$$

这步骤通常称为用正交变换化二次型为标准形（或者说把二次型化到主轴上去）. 其计算步骤如下：

1. 写出二次型矩阵 A；

2. 求出正交矩阵 T，使 $T'AT=D$ 为对角形；

3. 做变数替换 $X=TZ$ 后二次型化为标准形.

例 2 给定实二次型

$$f=2x_1x_2+2x_1x_3-2x_1x_4-2x_2x_3+2x_2x_4+2x_3x_4,$$

用正交变换将它化为标准形.

解 (i) 写出二次型矩阵

$$A=\begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix}.$$

(ii) 求正交矩阵 T，使 $T'AT=D$ 为对角形. 这已在例 1 中算出：

$$T=\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{12}} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & -\frac{1}{2} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & -\frac{1}{2} \\ 0 & 0 & \frac{3}{\sqrt{12}} & \frac{1}{2} \end{bmatrix}.$$

做变数替换 $X=TZ$：

$$\begin{cases} x_1=\frac{1}{\sqrt{2}}z_1+\frac{1}{\sqrt{6}}z_2-\frac{1}{\sqrt{12}}z_3+\frac{1}{2}z_4, \\ x_2=\frac{1}{\sqrt{2}}z_1-\frac{1}{\sqrt{6}}z_2+\frac{1}{\sqrt{12}}z_3-\frac{1}{2}z_4, \\ x_3=\frac{2}{\sqrt{6}}z_2+\frac{1}{\sqrt{12}}z_3-\frac{1}{2}z_4, \\ x_4=\frac{3}{\sqrt{12}}z_3+\frac{1}{2}z_4. \end{cases}$$

二次型化为标准形

$$z_1^2+z_2^2+z_3^2-3z_4^2.$$

习 题 三

1. 设 V 为 n 维欧氏空间，$\boldsymbol{A}$ 与 $\boldsymbol{A}^*$ 为 V 内两个线性变换. 如果

对任意 $\alpha,\beta\in V$ 有

$$(\boldsymbol{A}\alpha,\beta)=(\alpha,\boldsymbol{A}^{*}\beta),$$

则称 $\boldsymbol{A}^{*}$ 为 $\boldsymbol{A}$ 的**共轭变换**. 证明：$\boldsymbol{A}$ 与 $\boldsymbol{A}^{*}$ 在 V 的任一组标准正交基下的矩阵互为转置.

2. 续上题. 证明：对 V 内每个线性变换 $\boldsymbol{A}$，其共轭变换是存在且唯一的，而且 $(\boldsymbol{A}^{*})^{*}=\boldsymbol{A}$. 证明 $\boldsymbol{A}$ 是对称变换的充分必要条件是 $\boldsymbol{A}^{*}=\boldsymbol{A}$.

3. 证明：对 n 维欧氏空间 V 内任一线性变换 $\boldsymbol{A}$，$\boldsymbol{A}+\boldsymbol{A}^{*}$ 是一个对称变换.

4. 设 $\boldsymbol{A}$ 是 n 维欧氏空间 V 中的一个线性变换，如果 $\boldsymbol{A}^{*}=-\boldsymbol{A}$，即对任意 $\alpha,\beta\in V$，有

$$(\boldsymbol{A}\alpha,\beta)=-(\alpha,\boldsymbol{A}\beta),$$

则称 $\boldsymbol{A}$ 是一个**反对称变换**. 证明：

(1) $\boldsymbol{A}$ 为反对称变换的充分必要条件是：$\boldsymbol{A}$ 在某一组标准正交基下的矩阵是反对称矩阵；

(2) 如果 M 是反对称变换 $\boldsymbol{A}$ 的不变子空间，则 M 的正交补 $M^{\perp}$ 也是 $\boldsymbol{A}$ 的不变子空间.

5. 求正交矩阵 T，使 $T'AT$ 成对角形.

(1) $A=\begin{pmatrix}2&-2&0\\-2&1&-2\\0&-2&0\end{pmatrix}$；

(2) $A=\begin{pmatrix}2&2&-2\\2&5&-4\\-2&-4&5\end{pmatrix}$；

(3) $A=\begin{pmatrix}0&0&4&1\\0&0&1&4\\4&1&0&0\\1&4&0&0\end{pmatrix}$；

(4) $A=\begin{pmatrix}-1 & -3 & 3 & -3\\ -3 & -1 & -3 & 3\\ 3 & -3 & -1 & -3\\ -3 & 3 & -3 & -1\end{pmatrix}$;

(5) $A=\begin{pmatrix}1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\end{pmatrix}$.

6. 用正交变换化下列二次型成标准形.

(1) $f=x_1^2+2x_2^2+3x_3^2-4x_1x_2-4x_2x_3$;

(2) $f=x_1^2-2x_2^2-2x_3^2-4x_1x_2+4x_1x_3+8x_2x_3$;

(3) $f=2x_1x_2+2x_3x_4$;

(4) $f=x_1^2+x_2^2+x_3^2+x_4{}^2-2x_1x_2+6x_1x_3-4x_1x_4$
$-4x_2x_3+6x_2x_4-2x_3x_4$.

7. 设 A 是 n 阶实对称矩阵，证明：A 正定的充分必要条件是，A 的特征多项式的根全大于零.

8. 设 A,B 都是 n 阶实对称矩阵，证明：存在正交矩阵 T，使 $T^{-1}AT=B$ 的充分必要条件是，A,B 的特征多项式的根相同，且每个根的重数也相同.

9. 设 A,B 是 n 阶实对称矩阵，A 正定，证明：存在一可逆矩阵 T，使 $T'AT$ 和 $T'BT$ 同时成对角形.

10. 设 A 为正定矩阵，B 为实数矩阵.

(1) 证明：对于任意正整数 k，A^k 也正定.

(2) 如果对于某一正整数 r 有 $A^rB=BA^r$，证明：

$$AB=BA.$$

*第八章　酉空间与满秩双线性度量空间

在这一章中，我们研究实数域和复数域上线性空间中内积的更一般的定义方法.

§1　酉　空　间

在有了欧几里得空间的知识之后，很自然会提出这样的问题：能不能在复数域上的线性空间内设法定义内积，使它具有与欧氏空间相类似的性质呢？回答是肯定的，本节就解决这个问题.

酉空间的基本概念

定义　设 V 是复数域 $\mathscr{C}$ 上的线性空间. 如果给定一个法则，使 V 内任意两个向量 α,β 都按照这个法则对应于 $\mathscr{C}$ 内一个唯一确定的数，记作 (α,β)，且满足：

(i) 对任意 $k_1,k_2\in\mathscr{C},\alpha_1,\alpha_2,\beta\in V$，有

$$(k_1\alpha_1+k_2\alpha_2,\beta)=k_1(\alpha_1,\beta)+k_2(\alpha_2,\beta);$$

(ii) $(\alpha,\beta)=\overline{(\beta,\alpha)}$（取复数共轭），因此，对任意 $\alpha\in V$，(α,α) 都是实数；

(iii) 对任意 $\alpha\in V$，$(\alpha,\alpha)\geqslant 0$，且 $(\alpha,\alpha)=0\Longleftrightarrow\alpha=0$.

则称二元函数 (α,β) 为 V 内向量 α,β 的**内积**. 定义了这种内积的 $\mathscr{C}$ 上线性空间称为**酉空间**.

从内积的性质(i)与(ii)可得：对任意 $l_1,l_2\in\mathscr{C},\alpha,\beta_1,\beta_2\in V$，有

$$(\alpha,l_1\beta_1+l_2\beta_2)=\overline{(l_1\beta_1+l_2\beta_2,\alpha)}$$

$$
\begin{aligned}
&= \overline{l_1(\beta_1,\alpha) + l_2(\beta_2,\alpha)} \\
&= \bar{l}_1\,\overline{(\beta_1,\alpha)} + \bar{l}_2\,\overline{(\beta_2,\alpha)} \\
&= \bar{l}_1(\alpha,\beta_1) + \bar{l}_2(\alpha,\beta_2).
\end{aligned}
$$

由此可以看出，(α,β)对第二个变元 β 不是线性的，所以它不是双线性函数. 这是酉空间的内积与欧氏空间内积的一个重要区别. 其所以如此，是由于内积的性质(ii)与欧氏空间内积的相应性质有不同，而性质(ii)是保证(α,α)为实数的必需条件. 下面论述酉空间的一些基本概念.

一、向量的长度

因为(α,α)总是一个非负实数，我们定义

$$|\alpha| = \sqrt{(\alpha,\alpha)},$$

称为向量 α 的**模**或**长度**. $|\alpha|=1$ 时，α 称为**单位向量**. 我们有：对一切 $k\in\mathscr{C}$，

$$|k\alpha| = \sqrt{(k\alpha,k\alpha)} = \sqrt{k\bar{k}(\alpha,\alpha)} = |k|\cdot|\alpha|.$$

由此即知，对任一非零向量 α，$\dfrac{1}{|\alpha|}\alpha$ 为一单位向量，我们称之为 α 的**单位化**.

二、向量的正交性

在酉空间内两向量的内积(α,β)一般是一个复数，所以向量间没有夹角的概念，但却可以有正交的概念.

定义 一个酉空间 V 内两个向量 α 与 β 满足$(\alpha,\beta)=0$时称为互相**正交**，记作 $\alpha\perp\beta$.

注意$(\alpha,\beta)=0$时自然有

$$(\beta,\alpha) = \overline{(\alpha,\beta)} = 0.$$

另外，显然零向量与任意向量都正交.

三、内积的存在性

在有了酉空间内积的定义之后，自然会产生这样一个问题：满足条件(i)～(iii)的二元函数(α,β)是否存在呢？我们现在对有

限维线性空间来回答这一问题.

设 V 是 $\mathscr{C}$ 上的 n 维线性空间，在 V 内任取一组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n.$$

又设

$$\alpha = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n,$$

$$\beta = y_1\varepsilon_1 + y_2\varepsilon_2 + \cdots + y_n\varepsilon_n.$$

在 V 内定义二元函数如下：

$$(\alpha,\beta) = x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n}.$$

不难验证，这个二元函数即满足内积条件(i)～(iii). 显然，这时有

$$(\varepsilon_i,\varepsilon_j) = \delta_{ij} \quad (i,j = 1,2,\cdots,n).$$

四、酉空间的标准正交基

上一段的分析还给了我们这样一个启示，即在酉空间内可以有类似于欧氏空间中的标准正交基的概念. 为了引进这一重要概念，我们先指出一个简单的事实.

命题 1.1 酉空间 V 内两两正交的非零向量 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 所组成的向量组线性无关.

这个命题的证明与欧氏空间中的相应命题的证明相同，留给读者作为练习.

定义 在 n 维酉空间 V 内 n 个两两正交的单位向量组成的向量组称为 V 的一组**标准正交基**.

根据这个定义，V 内 n 个向量 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是一组标准正交基，等价于

$$(\varepsilon_i,\varepsilon_j) = \delta_{ij} \quad (i,j = 1,2,\cdots,n).$$

如果 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是一组标准正交基，设

$$\alpha = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n,$$

$$\beta = y_1\varepsilon_1 + y_2\varepsilon_2 + \cdots + y_n\varepsilon_n.$$

则

$$
\begin{aligned}
(\alpha,\beta) &= \left(\sum_{i=1}^{n} x_i\varepsilon_i, \sum_{j=1}^{n} y_j\varepsilon_j\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\overline{y}_j(\varepsilon_i,\varepsilon_j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \delta_{ij}x_i\overline{y}_j = \sum_{i=1}^{n} x_i\overline{y}_i.
\end{aligned}
$$

这就是在标准正交基下内积的表达式.它与欧氏空间中内积在标准正交基下的表达形式相似,只是现在第二个向量 β 的坐标要取复共轭.

五、标准正交基的求法

在酉空间内有与欧氏空间相同的施密特正交化方法.给定酉空间内一个线性无关向量组

$$\alpha_1,\alpha_2,\cdots,\alpha_s.$$

令

$$
\begin{aligned}
\varepsilon_1 &= \alpha_1, \\
\varepsilon_2 &= \alpha_2 - \frac{(\alpha_2,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1, \\
&\cdots\cdots \\
\varepsilon_{i+1} &= \alpha_{i+1} - \frac{(\alpha_{i+1},\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1 - \frac{(\alpha_{i+1},\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\varepsilon_2 - \cdots - \frac{(\alpha_{i+1},\varepsilon_i)}{(\varepsilon_i,\varepsilon_i)}\varepsilon_i, \\
&\cdots\cdots \\
\varepsilon_s &= \alpha_s - \frac{(\alpha_s,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1 - \frac{(\alpha_s,\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\varepsilon_2 - \cdots - \frac{(\alpha_s,\varepsilon_{s-1})}{(\varepsilon_{s-1},\varepsilon_{s-1})}\varepsilon_{s-1}.
\end{aligned}
$$

那么,同样有如下两条性质:

(i) $L(\varepsilon_1,\cdots,\varepsilon_i)=L(\alpha_1,\cdots,\alpha_i)\ (i=1,2,\cdots,s)$;

(ii) $(\varepsilon_i,\varepsilon_j)=0\ \ (i\neq j)$.

利用施密特正交化方法把一个有限维酉空间的一组基正交化后再单位化,就得到它的一组标准正交基.

六、标准正交基间的过渡矩阵

给定复数域上的一个 n 阶方阵 $A=(a_{ij})$. 令

$$\overline{A}=\begin{pmatrix}\overline{a}_{11} & \overline{a}_{12} & \cdots & \overline{a}_{1n}\\ \overline{a}_{21} & \overline{a}_{22} & \cdots & \overline{a}_{2n}\\ \vdots & \vdots & & \vdots\\ \overline{a}_{n1} & \overline{a}_{n2} & \cdots & \overline{a}_{nn}\end{pmatrix}$$

(即将 A 的每个元素取复共轭).

定义 设 U 是一个 n 阶可逆复矩阵. 如果 $\overline{U}'=U^{-1}$,则称 U 是一个**酉矩阵**.

如果把实数矩阵也看成一个复矩阵,它的每个元素取复共轭后没有变化. 由此可知,正交矩阵当作复矩阵看时就是酉矩阵. 所以,酉矩阵是正交矩阵的推广. 在欧氏空间中两组标准正交基之间的过渡矩阵是正交矩阵,对于酉空间,我们有类似的结果.

命题 1.2 设 V 是一个 n 维酉空间;$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组标准正交基;U 是一个 n 阶复方阵. 令

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)U,$$

则 $\eta_1,\eta_2,\cdots,\eta_n$ 是标准正交基的充分必要条件是:U 是一个酉矩阵.

证 (i) 必要性. 若 $\eta_1,\eta_2,\cdots,\eta_n$ 是标准正交基,则

$$(\eta_i,\eta_j)=\delta_{ij}\quad(i,j=1,2,\cdots,n).$$

设 $U=(u_{ij})$,U 的第 j 个列向量为 η_j 在 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的坐标,而 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 为标准正交基,故

$$(\eta_i,\eta_j)=u_{1i}\overline{u}_{1j}+u_{2i}\overline{u}_{2j}+\cdots+u_{ni}\overline{u}_{nj}=\delta_{ij}.$$

这表示 $\overline{U}'U=E$,即 $\overline{U}'=U^{-1}$,故 U 为酉矩阵.

(ii) 充分性. 若 U 为酉矩阵,则 $\overline{U}'U=E$. 于是有

$$(\eta_i,\eta_j)=u_{1i}\overline{u}_{1j}+u_{2i}\overline{u}_{2j}+\cdots+u_{ni}\overline{u}_{nj}=\delta_{ij}.$$

故 $\eta_1,\eta_2,\cdots,\eta_n$ 为 V 的一组标准正交基. ▌

酉 变 换

定义 设 U 是酉空间 V 内的一个线性变换,满足

$$(\boldsymbol{U}\alpha,\boldsymbol{U}\beta)=(\alpha,\beta)\quad(\text{对一切 } \alpha,\beta\in V),$$

则称 $\boldsymbol{U}$ 是一个**酉变换**.

酉变换不改变向量的内积，所以它应当与欧氏空间中的正交变换有类似的性质. 我们把这些性质概括为如下的命题.

命题 1.3 设 $\boldsymbol{U}$ 是 n 维酉空间 V 内的一个线性变换，则下列命题等价：

（i）$\boldsymbol{U}$ 是一个酉变换；

（ii）对任意 $\alpha\in V$，有 $|\boldsymbol{U}\alpha|=|\alpha|$；

（iii）$\boldsymbol{U}$ 把标准正交基变为标准正交基；

（iv）$\boldsymbol{U}$ 在标准正交基下的矩阵是酉矩阵.

证 类似于第七章命题 2.1 的证明，采用轮转证法.

（i）⇒（ii）. 显然.

（ii）⇒（iii）. 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组标准正交基. 由假设知 $|\boldsymbol{U}\varepsilon_i|=|\varepsilon_i|=1$，故只要证 $(\boldsymbol{U}\varepsilon_i,\boldsymbol{U}\varepsilon_j)=0\ (i\neq j)$. 对任意 $\alpha\in V$，有

$$(\boldsymbol{U}\alpha,\boldsymbol{U}\alpha)=|\boldsymbol{U}\alpha|^2=|\alpha|^2=(\alpha,\alpha).$$

以 $\alpha=k\varepsilon_i+\varepsilon_j$ 代入上式，展开后消去两边相等的项，得

$$k(\boldsymbol{U}\varepsilon_i,\boldsymbol{U}\varepsilon_j)+\bar{k}(\boldsymbol{U}\varepsilon_j,\boldsymbol{U}\varepsilon_i)=k(\varepsilon_i,\varepsilon_j)+\bar{k}(\varepsilon_j,\varepsilon_i)=0.$$

取 $k=1$ 及 i，得

$$(\boldsymbol{U}\varepsilon_i,\boldsymbol{U}\varepsilon_j)+(\boldsymbol{U}\varepsilon_j,\boldsymbol{U}\varepsilon_i)=0$$

及

$$\mathrm{i}(\boldsymbol{U}\varepsilon_i,\boldsymbol{U}\varepsilon_j)-\mathrm{i}(\boldsymbol{U}\varepsilon_j,\boldsymbol{U}\varepsilon_i)=0.$$

由此易知 $(\boldsymbol{U}\varepsilon_i,\boldsymbol{U}\varepsilon_j)=0$.

（iii）⇒（iv）. 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组标准正交基，$\boldsymbol{U}$ 在此基下的矩阵为 U. 矩阵 U 即是由基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 到基 $\boldsymbol{U}\varepsilon_1,\boldsymbol{U}\varepsilon_2,\cdots,\boldsymbol{U}\varepsilon_n$ 的过渡矩阵. 由假设，$\boldsymbol{U}\varepsilon_1,\boldsymbol{U}\varepsilon_2,\cdots,\boldsymbol{U}\varepsilon_n$ 也 V 的一组标准正交基. 根据命题 1.2，即知 U 是酉矩阵.

（iv）⇒（i）. 设 $\boldsymbol{U}$ 在标准正交基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵 U 是酉矩阵. 由命题 1.2 知，$\boldsymbol{U}\varepsilon_1,\boldsymbol{U}\varepsilon_2,\cdots,\boldsymbol{U}\varepsilon_n$ 也是 V 的标准正交基. 对任意的 $\alpha,\beta\in V$，设

$$\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n,$$

$$\beta = y_1\varepsilon_1 + y_2\varepsilon_2 + \cdots + y_n\varepsilon_n.$$

则
$$U\alpha = x_1U\varepsilon_1 + x_2U\varepsilon_2 + \cdots + x_nU\varepsilon_n,$$
$$U\beta = y_1U\varepsilon_1 + y_2U\varepsilon_2 + \cdots + y_nU\varepsilon_n.$$

由内积在标准正交基下的表达式，有
$$(U\alpha,U\beta) = x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n} = (\alpha,\beta),$$

即U是酉变换. ▌

命题 1.4 设V是一个n维酉空间. 令$U(n)$表示V内全体酉变换所成的集合，则有

(i) $E\in U(n)$;

(ii) 若$U_1,U_2\in U(n)$，则$U_1U_2\in U(n)$;

(iii) 若$U\in U(n)$，则U可逆，且$U^{-1}\in U(n)$.

证明留给读者作为练习.

与第七章§2中的$O(n)$一样，$U(n)$也是一个群，称为n阶**酉群**，它也有广泛的应用.

下面，我们介绍酉空间中一个子空间的正交补的概念.

定义 设V是n维酉空间，M是V的子空间. 令
$$M^{\perp} = \{\alpha \in V \mid (\alpha,m) = 0，对一切\ m \in M\}.$$

称$M^{\perp}$为M的**正交补**.

容易验证：$M^{\perp}$是V的子空间. 我们有

命题 1.5 $V=M\oplus M^{\perp}$.

证 在酉空间内，$(\alpha,\alpha)=0$等价于$\alpha=0$. 因此，$M\cap M^{\perp}=\{0\}$. 即$M+M^{\perp}$是直和. 如果在M中取一组标准正交基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r$，扩充成$V$的一组标准正交基(根据施密特正交化方法，这总是可以办到的)
$$\varepsilon_1,\cdots,\varepsilon_r,\varepsilon_{r+1},\cdots,\varepsilon_n.$$

任给$\alpha\in V$，有
$$\alpha = (k_1\varepsilon_1 + \cdots + k_r\varepsilon_r) + (k_{r+1}\varepsilon_{r+1} + \cdots + k_n\varepsilon_n),$$

其中
$$k_1\varepsilon_1 + \cdots + k_r\varepsilon_r \in M;\ k_{r+1}\varepsilon_{r+1} + \cdots + k_n\varepsilon_n \in M^{\perp}.$$

故 $M+M^{\perp}=V$，即 $V=M\oplus M^{\perp}$． ▌

习　题　一

1．在 $\mathscr{C}[X]_n$ 中定义二元函数如下

$$(f,g)=\sum_{k=1}^{n}f(k)\overline{g(k)}.$$

证明它满足内积条件(i)～(iii)，从而 $\mathscr{C}[X]_n$ 关于此内积成一酉空间.

2．在 $\mathscr{C}^n$ 中定义二元函数如下：若

$$\alpha=(a_1,a_2,\cdots,a_n);\ \beta=(b_1,b_2,\cdots,b_n),$$

则令

$$(\alpha,\beta)=a_1\overline{b}_1+a_2\overline{b}_2+\cdots+a_n\overline{b}_n.$$

证明它满足内积条件(i)～(iii)，从而 $\mathscr{C}^n$ 关于此内积成一酉空间. 证明：一个 n 阶复方阵 U 是酉矩阵的充分必要条件是：它的行(或列)向量组构成酉空间 $\mathscr{C}^n$ 的一组标准正交基.

3．在酉空间 $\mathscr{C}^n$ 中给定如下一组基

$$\alpha_1=(\mathrm{i},-1,\mathrm{i});\ \alpha_2=(1,0,\mathrm{i});\ \alpha_3=(1,1,1).$$

利用施密特正交化方法把它正交化后再单位化，求出 $\mathscr{C}^3$ 的一组标准正交基.

4．在题 1 的酉空间 $\mathscr{C}[X]_n$ 中，取 $n=3$，求出它的一组标准正交基.

5．证明：酉变换的特征值的模等于 1.

6．证明酉空间的哥西-布尼雅可夫斯基不等式

$$|(\alpha,\beta)|\leqslant|\alpha|\cdot|\beta|,$$

且等号成立的充分必要条件是：α 与 β 线性相关.

7．在酉空间中证明不等式

$$|\alpha+\beta|\leqslant|\alpha|+|\beta|.$$

8．在酉空间中定义两向量 α,β 的距离为

$$d(\alpha,\beta)=|\alpha-\beta|.$$

证明：

(1) $d(\alpha,\beta)\geqslant 0$，且 $d(\alpha,\beta)=0\Longleftrightarrow\alpha=\beta$；

(2) $d(\alpha,\beta)=d(\beta,\alpha)$；

(3) $d(\alpha,\gamma)\leqslant d(\alpha,\beta)+d(\beta,\gamma)$.

9. 设 $\boldsymbol{U}$ 是 n 维酉空间 V 内的一个可逆线性变换，证明：如果对任意 $\alpha,\beta\in V$，有

$$(\boldsymbol{U}\alpha,\beta)=(\alpha,\boldsymbol{U}^{-1}\beta),$$

则 $\boldsymbol{U}$ 是一个酉变换.

10. 设 $\boldsymbol{U}$ 为 n 维酉空间 V 内的一个酉变换，其全部特征值设为 $\lambda_1,\lambda_2,\cdots,\lambda_n$. 证明：$\overline{\lambda_1},\overline{\lambda_2},\cdots,\overline{\lambda_n}$ 是 $\boldsymbol{U}^{-1}$的全部特征值.

11. 将一个复方阵 U 分解为实部和虚部

$$U=P+\mathrm{i}Q$$

（其中 P,Q 为实 n 阶方阵）. 证明 U 为酉矩阵的充分必要条件是：$P'Q$ 对称，且 $P'P+Q'Q=E$.

*12. 证明下面矩阵是酉矩阵

$$U=\frac{1}{\sqrt{n}}\begin{pmatrix}1 & 1 & \cdots & 1\\ 1 & \omega & \cdots & \omega^{n-1}\\ \vdots & \vdots & & \vdots\\ 1 & \omega^{n-1} & \cdots & \omega^{(n-1)^2}\end{pmatrix}\quad(\omega=\mathrm{e}^{\frac{2\pi\mathrm{i}}{n}}).$$

13. 证明任一个二阶酉矩阵 U 可分解为

$$U=\begin{pmatrix}\mathrm{e}^{\mathrm{i}\theta_1} & 0\\ 0 & \mathrm{e}^{\mathrm{i}\theta_2}\end{pmatrix}\begin{pmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{pmatrix}\begin{pmatrix}\mathrm{e}^{\mathrm{i}\theta_3} & 0\\ 0 & \mathrm{e}^{\mathrm{i}\theta_4}\end{pmatrix},$$

其中 $\theta_1,\theta_2,\theta_3,\theta_4,\varphi$ 为实数.

§2 正规变换与厄米特变换

在这一节里，我们将把欧氏空间中的对称变换推广到酉空间，得到大致相平行的结果. 但在这里，我们将从更一般的角度来观察

问题.

首先给出酉空间中一组基的度量矩阵的概念.

定义 在 n 维酉空间 V 内取定一组基

$$\eta_1,\eta_2,\cdots,\eta_n.$$

令

$$(\eta_i,\eta_j)=g_{ij}\quad (i,j=1,2,\cdots,n).$$

称 n 阶方阵 $G=(g_{ij})$ 为基 $\eta_1,\eta_2,\cdots,\eta_n$ 的**度量矩阵**.

命题 2.1 n 维酉空间 V 内任意一组基 $\eta_1,\eta_2,\cdots,\eta_n$ 的度量矩阵 $G=(g_{ij})$ 都是可逆的.

证 在 V 内取一组标准正交基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,设两组基间的过渡矩阵为 $T=(t_{ij})$,即

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T.$$

那么,有

$$\eta_i=t_{1i}\varepsilon_1+t_{2i}\varepsilon_2+\cdots+t_{ni}\varepsilon_n,$$

$$g_{ij}=(\eta_i,\eta_j)=t_{1i}\bar{t}_{1j}+t_{2i}\bar{t}_{2j}+\cdots+t_{ni}\bar{t}_{nj}.$$

写成矩阵形式,就是 $G=T'\overline{T}$. 因而,

$$|G|=|T'\overline{T}|=|T'|\,|\overline{T}|=|T|\cdot|\overline{T}|\neq 0.$$

故 G 可逆. ▌

定义 设 $\boldsymbol{A}$ 是 n 维酉空间 V 内的一个线性变换. 如果 V 内一个线性变换 $\boldsymbol{A}^*$ 满足如下条件:对一切 $\alpha,\beta\in V$,有

$$(\boldsymbol{A}\alpha,\beta)=(\alpha,\boldsymbol{A}^*\beta),$$

则称 $\boldsymbol{A}^*$ 为 $\boldsymbol{A}$ 的**共轭变换**.

现在在 V 内取一组基

$$\eta_1,\eta_2,\cdots,\eta_n.$$

设其度量矩阵为 $G=(g_{ij})$. 而 $\boldsymbol{A}$ 与 $\boldsymbol{A}^*$ 在这组基下的矩阵分别为 $A=(a_{ij})$,$B=(b_{ij})$. $\boldsymbol{A}^*$ 是 $\boldsymbol{A}$ 的共轭变换,它事实上等价于

$$(\boldsymbol{A}\eta_i,\eta_j)=(\eta_i,\boldsymbol{A}^*\eta_j)\quad (i,j=1,2,\cdots,n).$$

具体写出,就是

$$\sum_{k=1}^{n} a_{ik} g_{kj} = \sum_{k=1}^{n} g_{ik} \overline{b}_{kj}.$$

写成矩阵形式，就是

$$A'G = G\overline{B}.$$

因为度量矩阵 G 是可逆的，所以上式等价于

$$\overline{B} = G^{-1}A'G,$$

$$B = \overline{G}^{-1}\overline{A}'\overline{G}. \qquad (1)$$

如果 $\boldsymbol{A}^*$ 是 $\boldsymbol{A}$ 的共轭变换，则 B 满足(1)式. 反之，若 B 满足(1)式，$\boldsymbol{A}^*$ 就是 $\boldsymbol{A}$ 的共轭变换. 由此可知：V 内任一线性变换的共轭变换都存在而且唯一. 如果 $\eta_1,\eta_2,\cdots,\eta_n$ 是一组标准正交基，则 $G=E$，(1)式变成 $B=\overline{A}'$. 即 $\boldsymbol{A}$ 的共轭变换 $\boldsymbol{A}^*$ 在标准正交基下的矩阵恰为 $\boldsymbol{A}$ 的矩阵 A 取复共轭以后再转置. 因此，一个酉变换 $\boldsymbol{U}$ 的共轭变换即为其逆变换：$\boldsymbol{U}^*=\boldsymbol{U}^{-1}$(因为 $\boldsymbol{U}$ 在标准正交基下的矩阵为酉矩阵，取了复共轭再转置时恰为其逆矩阵). 由此可知，对于酉变换，有 $\boldsymbol{U}\boldsymbol{U}^*=\boldsymbol{U}^*\boldsymbol{U}=\boldsymbol{E}$.

不难验证有如下关系式：

(i) $\boldsymbol{E}^*=\boldsymbol{E}$;

(ii) $(\boldsymbol{A}^*)^*=\boldsymbol{A}$;

(iii) $(\lambda\boldsymbol{A})^*=\overline{\lambda}\boldsymbol{A}^*$;

(iv) $(\boldsymbol{A}+\boldsymbol{B})^*=\boldsymbol{A}^*+\boldsymbol{B}^*$;

(v) $(\boldsymbol{A}\boldsymbol{B})^*=\boldsymbol{B}^*\boldsymbol{A}^*$.

例如，我们证明(iii)与(v).

(iii)：对任意 $\alpha,\beta\in V$，有

$$((\lambda\boldsymbol{A})\alpha,\beta) = \lambda(\boldsymbol{A}\alpha,\beta) = \lambda(\alpha,\boldsymbol{A}^*\beta) = (\alpha,(\overline{\lambda}\boldsymbol{A}^*)\beta).$$

由于 $\lambda\boldsymbol{A}$ 的共轭变换存在且唯一，故由上面的等式即可推断

$$(\lambda\boldsymbol{A})^* = \overline{\lambda}\boldsymbol{A}^*.$$

(v)：对任意 $\alpha,\beta\in V$，有

$$((\boldsymbol{A}\boldsymbol{B})\alpha,\beta) = (\boldsymbol{B}\alpha,\boldsymbol{A}^*\beta) = (\alpha,(\boldsymbol{B}^*\boldsymbol{A}^*)\beta).$$

因为 $\boldsymbol{AB}$ 的共轭变换存在且唯一,故由上面的等式即可推断

$$(\boldsymbol{AB})^* = \boldsymbol{B}^*\boldsymbol{A}^*.$$

定义 酉空间 V 内一个线性变换 $\boldsymbol{A}$ 如与其共轭变换可交换:$\boldsymbol{AA}^* = \boldsymbol{A}^*\boldsymbol{A}$,则称为一个**正规变换**.

根据前面的分析可知,酉变换是一种正规变换.

下面我们阐述正规变换的几个重要性质.

命题 2.2 设 $\boldsymbol{A}$ 是酉空间 V 内的一个正规变换,而 λ 是 $\boldsymbol{A}$ 的一个特征值,其对应特征向量为 ξ. 那么,ξ 是 $\boldsymbol{A}^*$ 的属于特征值 $\bar{\lambda}$ 的特征向量.

证 按假设,有 $\boldsymbol{A}\xi - \lambda\xi = (\boldsymbol{A} - \lambda\boldsymbol{E})\xi = 0$. 于是

$$\begin{aligned}
0 &= ((\boldsymbol{A} - \lambda\boldsymbol{E})\xi, (\boldsymbol{A} - \lambda\boldsymbol{E})\xi) \\
&= (\xi, (\boldsymbol{A} - \lambda\boldsymbol{E})^*(\boldsymbol{A} - \lambda\boldsymbol{E})\xi) \\
&= (\xi, (\boldsymbol{A} - \bar{\lambda}\boldsymbol{E}^*)(\boldsymbol{A} - \lambda\boldsymbol{E})\xi) \\
&= (\xi, (\boldsymbol{A} - \lambda\boldsymbol{E})(\boldsymbol{A}^* - \bar{\lambda}\boldsymbol{E})\xi) \\
&= \overline{((\boldsymbol{A} - \lambda\boldsymbol{E})(\boldsymbol{A}^* - \lambda\boldsymbol{E})\xi, \xi)} \\
&= \overline{((\boldsymbol{A}^* - \lambda\boldsymbol{E})\xi, (\boldsymbol{A} - \lambda\boldsymbol{E})^*\xi)} \\
&= ((\boldsymbol{A}^* - \bar{\lambda}\boldsymbol{E})\xi, (\boldsymbol{A}^* - \bar{\lambda}\boldsymbol{E})\xi).
\end{aligned}$$

由此即得 $(\boldsymbol{A}^* - \bar{\lambda}\boldsymbol{E})\xi = 0$. 于是 $\boldsymbol{A}^*\xi = \bar{\lambda}\xi$. ▌

命题 2.3 设 $\boldsymbol{A}$ 是酉空间 V 内的一个正规变换,则 $\boldsymbol{A}$ 的属于不同特征值的特征向量互相正交.

证 设 λ,μ 是 $\boldsymbol{A}$ 的两个互不相同的特征值,ξ,η 是分别属于 λ,μ 的特征值向量. 由命题 2.2 知:$\boldsymbol{A}^*\eta = \bar{\mu}\eta$. 于是有

$$\begin{aligned}
\lambda(\zeta,\eta) &= (\lambda\xi,\eta) = (\boldsymbol{A}\xi,\eta) = (\xi,\boldsymbol{A}^*\eta) \\
&= (\xi,\bar{\mu}\eta) = \mu(\xi,\eta).
\end{aligned}$$

移项,得
$$(\lambda - \mu)(\xi,\eta) = 0.$$

因 $\lambda - \mu \neq 0$,故 $(\xi,\eta) = 0$. ▌

下面是关于正规变换的基本定理.

定理 1 设 $\boldsymbol{A}$ 是 n 维酉空间 V 内的一个正规变换,则在 V 内

存在一组标准正交基，使 $\boldsymbol{A}$ 在这组基下的矩阵成对角形.

证 对 V 的维数 n 作数学归纳法.

$n=1$ 时命题显然成立. 设对 $n-1$ 维酉空间命题成立，证明对 n 维酉空间命题也成立.

设 λ_1 是 $\boldsymbol{A}$ 的一个特征值，η_1 是一个对应的特征向量，$|\eta_1|=1$. 命

$$M=\{\alpha\in V\mid(\eta_1,\alpha)=0\}.$$

显然，M 是 V 的一个子空间. 根据施密特正交化方法，η_1 可以扩充成 V 的一组标准正交基

$$\eta_1,\eta_2',\cdots,\eta_n'.$$

显然，$M=L(\eta_2',\cdots,\eta_n')$，故 $\dim M=n-1$. 我们证明 M 是 $\boldsymbol{A}$ 和 $\boldsymbol{A}^*$ 的不变子空间：设 $\alpha\in M$，由命题 2.2，有

$$(\boldsymbol{A}\alpha,\eta_1)=(\alpha,\boldsymbol{A}^*\eta_1)=(\alpha,\bar{\lambda}\eta_1)=\lambda(\alpha,\eta_1)=0.$$

于是 $\boldsymbol{A}\alpha\in M$，所以 M 是 $\boldsymbol{A}$ 的不变子空间. 又因为 $(\boldsymbol{A}^*)^*=\boldsymbol{A}$，故又有

$$\begin{aligned}(\boldsymbol{A}^*\alpha,\eta_1)&=(\alpha,(\boldsymbol{A}^*)^*\eta_1)=(\alpha,\boldsymbol{A}\eta_1)\\&=(\alpha,\lambda_1\eta_1)=\bar{\lambda}_1(\alpha,\eta_1)=0.\end{aligned}$$

于是 $\boldsymbol{A}^*\alpha\in M$，即 M 也是 $\boldsymbol{A}^*$ 的不变子空间. M 关于 V 的内积仍为一酉空间. $\boldsymbol{A}$ 在 M 内的限制 $\boldsymbol{A}|_M$ 的共轭变换恰为 $\boldsymbol{A}^*$ 在 M 内的限制 $\boldsymbol{A}^*|_M$，$\boldsymbol{A}$ 与 $\boldsymbol{A}^*$ 限制在 M 内当然还是交换的，因此，$\boldsymbol{A}$ 限制在 M 内仍为正规变换. 按归纳假设，在 M 内存在一组标准正交基 $\eta_2,\cdots,\eta_n$，使

$$\boldsymbol{A}\eta_i=\lambda_i\eta_i\quad(i=2,3,\cdots,n).$$

因为 $\eta_1,\eta_2,\cdots,\eta_n$ 满足关系式

$$(\eta_i,\eta_j)=\delta_{ij}\quad(i,j=1,2,\cdots,n),$$

故它们是 V 的一组标准正交基. 在这组基下 $\boldsymbol{A}$ 的矩阵成对角形：

$$\boldsymbol{A}\eta_i=\lambda_i\eta_i\quad(i=1,2,\cdots,n).\ \blacksquare$$

推论 设 $\boldsymbol{U}$ 是 n 维酉空间 V 内的一个酉变换，则在 V 内存在

一组标准正交基,使 $\boldsymbol{U}$ 在这组基下的矩阵成对角形.

这是因为 $\boldsymbol{U}$ 变换是一种正规变换的缘故.

下面我们再来研究另一类重要的正规变换,它可以看作是欧氏空间中的对称变换的自然推广.

定义 设 $\boldsymbol{A}$ 是酉空间 V 内一个线性变换,且 $\boldsymbol{A}^*=\boldsymbol{A}$. 则称 $\boldsymbol{A}$ 是一个**厄米特**(Hermite)**变换**.

厄米特变换显然是一个正规变换. 所以,前面关于正规变换所获得的结果对它都适用. 特别地,根据命题 2.2,如果 λ 是厄米特变换 $\boldsymbol{A}$ 的任一特征值,ξ 是其对应的特征向量,则

$$\bar{\lambda}\xi = \boldsymbol{A}^*\xi = \boldsymbol{A}\xi = \lambda\xi.$$

因 $\xi\neq 0$,故 $\bar{\lambda}=\lambda$. 因此,有

命题 2.4 厄米特变换的特征值都是实数.

综合定理 1 和命题 2.4,可以得到关于厄米特变换的如下重要结论:

定理 2 设 $\boldsymbol{A}$ 是 n 维酉空间 V 中的一个厄米特变换,则在 V 中存在一组标准正交基,使 $\boldsymbol{A}$ 在这组基下的矩阵是实对角矩阵.

在 n 维酉空间 V 中取一组标准正交基

$$\eta_1,\eta_2,\cdots,\eta_n.$$

设 $\boldsymbol{A}$ 是 V 内的一个厄米特变换,它在这组基下的矩阵为 A. 根据前面的(1)式可知,$\boldsymbol{A}^*$ 在这组基下的矩阵为 $\bar{A}'$,但 $\boldsymbol{A}^*=\boldsymbol{A}$,故必有 $\bar{A}'=A$.

定义 设 A 是一个 n 阶复方阵. 如果 $\bar{A}'=A$,则称 A 是一个**厄米特矩阵**.

显然,酉空间内一个线性变换 $\boldsymbol{A}$ 是厄米特变换的充分必要条件是:它在某一组标准正交基下的矩阵是厄米特矩阵. 反之,任一厄米特矩阵也可以看作一个酉空间中某个厄米特变换在一组标准正交基下的矩阵. 于是从定理 2 可得:

推论 设 A 是 n 阶厄米特矩阵,则存在一个 n 阶酉矩阵 U,使 $U^{-1}AU=\bar{U}'AU=D$ 是一个实对角矩阵.

现在把定理 2 及其推论应用到如下的厄米特二次型.

定义 n 个复变量 $x_1,x_2,\cdots,x_n$ 的二次齐次形式

$$f = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\bar{x}_i x_j \quad (a_{ij} = \bar{a}_{ji}) \tag{2}$$

称为一个**厄米特二次型**. 命

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

称 A 为 f **的矩阵**.

显然,一个厄米特二次型的矩阵是一个厄米特矩阵. 我们可以把厄米特二次型用矩阵形式表示如下:

$$f = \bar{X}'AX.$$

定理 3 对厄米特二次型(2),存在一个酉矩阵 U,使在酉线性变数替换 $X=UY$ 下它变为如下的标准形

$$d_1\bar{y}_1y_1 + d_2\bar{y}_2y_2 + \cdots + d_n\bar{y}_ny_n,$$

其中 $d_1,d_2,\cdots,d_n$ 均为实数,且除排列次序外,是被 f 唯一确定的.

证 f 的矩阵 A 是一个厄米特矩阵. 由定理 2 的推论,存在酉矩阵 U,使

$$\bar{U}'AU = D = \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}$$

为实对角矩阵. 令 $Y=U^{-1}X$,即 $X=UY$,代入

$$f = \bar{X}'AX = (\overline{UY})'A(UY) = \bar{Y}'(\bar{U}'AU)Y$$
$$= \bar{Y}'DY = d_1\bar{y}_1y_1 + d_2\bar{y}_2y_2 + \cdots + d_n\bar{y}_ny_n.$$

为了证明 $d_1,d_2,\cdots,d_n$ 的唯一性(差一个次序),我们先指出如下事实:如果对一切 $y_1,y_2,\cdots,y_n$,有

$$\overline{Y}'BY = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}\overline{y}_i y_j = \sum_{i=1}^{n} d_i \overline{y}_i y_i = \overline{Y}'DY,$$

则必定 $B=D$. 这是因为

（ i ） 令 $y_i=1, y_j=0(j\neq i$ 时)代入上式，得

$$b_{ii} = d_i \quad (i = 1,2,\cdots,n);$$

（ii） 对 $1\leqslant j<k\leqslant n$，令 $y_j=y_k=1, y_l=0(l\neq j,k$ 时)，代入上式，利用(i)中的结果，消去等式两边相等的项，得

$$b_{jk} + b_{kj} = 0. \tag{3}$$

又令 $y_j=\mathrm{i}$(此处 i 表虚数单位)，$y_k=1, y_l=0(l\neq j,k$ 时)，代入上式，两边消去公共因子 i，得

$$-b_{jk} + b_{kj} = 0. \tag{4}$$

把(3)，(4)两式相加，得 $b_{kj}=0$，代入(3)得 $b_{jk}=0$.

综合(i)与(ii)得 $B=D$.

现设 f 经酉线性变数替换 $X=UY$ 化为标准形

$$f= \overline{X}'AX \xlongequal{X=UY} \overline{Y}'(\overline{U}'AU)Y$$

$$= d_1\overline{y}_1y_1 + d_2\overline{y}_2y_2 + \cdots + d_n\overline{y}_ny_n = \overline{Y}'DY.$$

则 $D=\overline{U}'AU=U^{-1}AU$. 即 D 与 A 相似. 于是 $d_1,d_2,\cdots,d_n$ 恰为 A 的 n 个特征值，因而由 f 唯一确定(除差一个排列次序外). ▌

习 题 二

1. 对酉空间的共轭变换证明如下关系式：

（ i ） $\boldsymbol{E}^*=\boldsymbol{E}$;

（ii） $(\boldsymbol{A}^*)^*=\boldsymbol{A}$;

（iii） $(\boldsymbol{A}+\boldsymbol{B})^*=\boldsymbol{A}^*+\boldsymbol{B}^*$.

2. 设 $\boldsymbol{A}$ 是 n 维酉空间 V 内的一个线性变换，M 是 $\boldsymbol{A}$ 的不变子空间，证明：M 的正交补 $M^{\perp}$ 是 $\boldsymbol{A}^*$ 的不变子空间.

3. 设 $\boldsymbol{A}$ 是 n 维酉空间 V 内的一个线性变换. 如果存在一个复系数多项式 $f(\lambda)$，使 $\boldsymbol{A}=f(\boldsymbol{A}^*)$，证明在 V 内存在一组标准正

交基，使 $\boldsymbol{A}$ 在这组基下的矩阵成对角形.

4. 设 $\boldsymbol{A}$ 是 n 维酉空间 V 内的一个线性变换，$\boldsymbol{A}^* = -\boldsymbol{A}$. 证明：$\boldsymbol{A}$ 的非零特征值都是纯虚数.

5. 设 $\boldsymbol{A}$ 是 n 维酉空间 V 中的一个厄米特变换，证明：对任意 $\alpha \in V$，$(\boldsymbol{A}\alpha, \alpha)$ 是一个实数.

6. 设 $\boldsymbol{A}$ 是 n 维酉空间 V 中的一个厄米特变换. 如果对 V 中任一非零向量 α 都有

$$(\boldsymbol{A}\alpha, \alpha) > 0,$$

则称 $\boldsymbol{A}$ 为**正定厄米特变换**. 证明：一个厄米特变换 $\boldsymbol{A}$ 正定的充分必要条件是其特征值都大于零.

7. 证明：任一可逆厄米特变换 $\boldsymbol{A}$ 的平方 $\boldsymbol{A}^2$ 是正定厄米特变换.

8. 证明：任一可逆线性变换 $\boldsymbol{A}$ 与其共轭变换 $\boldsymbol{A}^*$ 的乘积 $\boldsymbol{A}\boldsymbol{A}^*$ 是正定厄米特变换.

9. 设 $\boldsymbol{A}$ 是 n 维酉空间 V 内的厄米特变换，$\lambda_1, \lambda_2, \cdots, \lambda_s$ 是 $\boldsymbol{A}$ 的全部互不相同的特征值，证明：

$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_s}.$$

10. 设 $\boldsymbol{A}$ 是 n 维酉空间 V 内的厄米特变换，λ_0 是它的一个 k 重特征值，证明 $\dim V_{\lambda_0} = k$.

11. 设 $\boldsymbol{A}$ 是 n 维酉空间中的一个正定厄米特变换，它在标准正交基下的矩阵 A 称为**正定厄米特矩阵**. 以 A 为矩阵的厄米特二次型 $\overline{X}'AX$ 称为**正定厄米特型**. 证明：任一正定厄米特型可用可逆线性变数替换 $X = TZ$ 化为

$$\bar{z}_1 z_1 + \bar{z}_2 z_2 + \cdots + \bar{z}_n z_n.$$

12. 设

$$f = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \bar{x}_i x_j \quad (a_{ij} = \bar{a}_{ji}),$$

$$g = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} \bar{x}_i x_j \quad (b_{ij} = \bar{b}_{ji}),$$

其中 f 是正定厄米特二次型. 证明:存在一个可逆线性变数替换 $X=TZ$,使 f 变为

$$\bar{z}_1z_1+\bar{z}_2z_2+\cdots+\bar{z}_nz_n;$$

而 g 变为

$$d_1\bar{z}_1z_1+d_2\bar{z}_2z_2+\cdots+d_n\bar{z}_nz_n.$$

§3 满秩对称双线性度量空间

在§1中所定义的酉空间,其内积不是双线性函数. 因为只有这样,才能保证对任一非零向量 α,(α,α)为正实数,因而才能有向量长度的概念(但夹角的概念却已经没有了). 这说明:对复数域上的线性空间,要保留某些类似于欧氏空间的度量性质,就要破坏内积的双线性这一性质. 反之,如果保持内积的双线性,那所得的空间的度量性质就与欧氏空间有很大的差别. 但尽管如此,引进这样的度量对研究许多问题仍然是很有用的. 本节和下一节就是要研究以某些双线性函数作为内积时,空间的一些基本度量性质.

准欧几里得空间

我们首先研究比较一般的双线性函数作为内积的实数域上线性空间.

定义 设 V 是实数域上的一个 n 维线性空间,$f(\alpha,\beta)$是 V 上一个满秩对称双线性函数. 定义 V 内两个向量 α,β 的**内积**为

$$(\alpha,\beta)=f(\alpha,\beta).$$

称具有这种内积的线性空间为**准欧几里得空间**,简称**准欧氏空间**.

准欧氏空间与欧氏空间的区别是:欧氏空间中的内积以正定矩阵作为其度量矩阵,而准欧氏空间没有这一要求. 因此,准欧氏空间是较欧氏空间更为普遍的概念.

现在我们介绍准欧氏空间的一些基本概念.

一、基的度量矩阵

在一个 n 维准欧氏空间内取一组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n.$$

令

$$a_{ij}=(\varepsilon_i,\varepsilon_j)\quad(i,j=1,2,\cdots,n),$$

则称 $A=(a_{ij})$是内积(α,β)在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的**度量矩阵**. 度量矩阵是一个实对称矩阵. 若设

$$\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n,$$
$$\beta=y_1\varepsilon_1+y_2\varepsilon_2+\cdots+y_n\varepsilon_n,$$

则

$$(\alpha,\beta)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_iy_j=X'AY.$$

二、规范基

因为内积在任一组基下的矩阵都是实对称矩阵，根据第六章§3 的定理 3，在 V 内可以找到这样的基 $\eta_1,\eta_2,\cdots,\eta_n$，使$(\alpha,\beta)$在这组基下的度量矩阵为

$$G=\begin{bmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&-1&&\\&&&&\ddots&\\&&&&&-1\end{bmatrix}\ (p\text{个}1).$$

如设

$$\alpha=x_1\eta_1+x_2\eta_2+\cdots+x_n\eta_n,$$
$$\beta=y_1\eta_1+y_2\eta_2+\cdots+y_n\eta_n,$$

则

$$(\alpha,\beta)=x_1y_1+\cdots+x_py_p-x_{p+1}y_{p+1}-\cdots-x_ny_n=X'GY.\tag{1}$$

$$(\alpha,\alpha)=x_1^2+\cdots+x_p^2-x_{p+1}^2-\cdots-x_n^2,$$

其中 p 是二次型函数(α,α)的正惯性指数，是唯一确定的. 这个二次型函数的符号差 $2p-n$ 称为该准欧氏空间 V 的**符号差**. 而基

$\eta_1, \eta_2, \cdots, \eta_n$ 则称为 V 的一组**规范基**. 在规范基下二次型函数 (α,α)成为规范形.

从(1)式可以看到，在准欧氏空间内，一向量 α 与自己的内积 (α,α)可能是负数，非零向量 α 与自己的内积(α,α)有可能是零. 因而，向量的长度和夹角的概念都没有了. 但是正交性的概念却保留着. 这就是说，V 内两个向量 α,β 如满足$(\alpha,\beta)=0$，则称 α 与 β 正交. 由于内积是对称的，所以正交性也具有对称性. 要注意的是：现在一个非零向量有可能和自己正交.

三、准欧氏空间中的正交变换

定义 设 $\boldsymbol{R}$ 是 n 维准欧氏空间 V 内的一个线性变换，如果对任意 $\alpha,\beta\in V$，都有

$$(\boldsymbol{R}\alpha,\boldsymbol{R}\beta) = (\alpha,\beta),$$

则称 $\boldsymbol{R}$ 是一个**正交变换**.

现在在 V 内取一组规范基

$$\eta_1,\eta_2,\cdots,\eta_n.$$

设 $\boldsymbol{R}$ 在这组基下的矩阵是 $R=(r_{ij})$. $\boldsymbol{R}$ 是正交变换，等价于

$$(\boldsymbol{R}\eta_i,\boldsymbol{R}\eta_j) = (\eta_i,\eta_j). \tag{2}$$

命 G 为基 $\eta_1,\eta_2,\cdots,\eta_n$ 的度量矩阵. 因

$$\boldsymbol{R}\eta_i = r_{1i}\eta_1 + r_{2i}\eta_2 + \cdots + r_{ni}\eta_n,$$

利用规范基下内积表达式(1)即知

$$(\boldsymbol{R}\eta_i,\boldsymbol{R}\eta_j) = (r_{1i} \quad r_{2i} \quad \cdots \quad r_{ni})G\begin{pmatrix} r_{1j} \\ r_{2j} \\ \vdots \\ r_{nj} \end{pmatrix} \quad (i,j=1,2,\cdots,n).$$

因而(2)式等价于如下矩阵关系式

$$R'GR = G. \tag{3}$$

由此即得：

命题 3.1 n 维准欧氏空间 V 内一个线性变换 $\boldsymbol{R}$ 是一个正交变换的充分必要条件是：它在 V 的某一组规范基下的矩阵 R 满足

(3)式.

命 $O_n(V)$ 表 n 维准欧氏空间 V 内全体正交变换所成的集合，我们有：

命题 3.2 $O_n(V)$ 有如下性质：

(i) $\boldsymbol{E}\in O_n(V)$；

(ii) 若 $\boldsymbol{R}_1,\boldsymbol{R}_2\in O_n(V)$，则 $\boldsymbol{R}_1\boldsymbol{R}_2\in O_n(V)$；

(iii) 若 $\boldsymbol{R}\in O_n(V)$，则 $\boldsymbol{R}$ 可逆，且 $\boldsymbol{R}^{-1}\in O_n(V)$.

证明留给读者作为练习.

复欧几里得空间

定义 设 V 是复数域 $\mathscr{C}$ 上的 n 维线性空间，$f(\alpha,\beta)$ 是 V 内一个满秩对称双线性函数. 规定 V 内两个向量 α,β 的内积为

$$(\alpha,\beta)=f(\alpha,\beta).$$

称具有这种内积的线性空间 V 为**复欧几里得空间**，简称**复欧氏空间**.

下面介绍复欧氏空间的一些基本概念.

一、正交性

设 α,β 是 V 内两个向量. 如果 $(\alpha,\beta)=0$，则称 α 与 β 正交. 因为内积是对称的，所以正交性也是对称的. 一个非零向量可能和自己正交：$(\alpha,\alpha)=0$. 但因为 $f(\alpha,\beta)$ 是满秩的，所以，如果一个向量 α 与 V 内所有向量正交，则 $\alpha=0$. 这一事实的证明留给读者作为练习.

二、基的度量矩阵

设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组基，令

$$g_{ij}=(\varepsilon_i,\varepsilon_j)\quad(i,j=1,2,\cdots,n).$$

称 $G=(g_{ij})$ 为这组基的度量矩阵. 显然，G 是复对称矩阵. 如设

$$\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n,$$

$$\beta=y_1\varepsilon_1+y_2\varepsilon_2+\cdots+y_n\varepsilon_n,$$

则

$$(\alpha,\beta)=\sum_{i=1}^{n}\sum_{j=1}^{n}g_{ij}x_iy_j=X'GY.$$

三、规范基

因为(α,β)是$\mathscr{C}$上线性空间内的满秩对称双线性函数，从第六章§3的定理2可知，在V内存在一组基$\eta_1,\eta_2,\cdots,\eta_n$，使内积在这组基下的度量矩阵为单位矩阵$E$，即

$$(\eta_i,\eta_j)=\delta_{ij}. \tag{4}$$

这类基称为**第一类规范基**. 如令

$$\alpha=x_1\eta_1+x_2\eta_2+\cdots+x_n\eta_n,$$
$$\beta=y_1\eta_1+y_2\eta_2+\cdots+y_n\eta_n,$$

则

$$(\alpha,\beta)=x_1y_1+x_2y_2+\cdots+x_ny_n.$$

这与欧氏空间中向量内积在标准正交基下的表达式相似，只是要注意现在$x_1,x_2,\cdots,x_n$与$y_1,y_2,\cdots,y_n$一般都是复数，因而(α,α)一般不是实数.

在复欧氏空间中常用的还有另一类基. 我们分两种情况对它进行讨论.

1. $\dim V=n=2m$.

此时令(命i表虚单位)

$$\left.\begin{aligned}\xi_j&=\frac{1}{\sqrt{2}}(\eta_j+\mathrm{i}\eta_{n-j+1})\\ \xi_{n-j+1}&=\frac{1}{\sqrt{2}}(\eta_j-\mathrm{i}\eta_{n-j+1})\end{aligned}\right\}\quad(j=1,2,\cdots,m).$$

显然，$\xi_1,\xi_2,\cdots,\xi_n$与$\eta_1,\eta_2,\cdots,\eta_n$等价，所以它也是$V$的一组基. 我们来求这组基的度量矩阵. 注意$\xi_j$和$\xi_{n-j+1}$的表达式中仅出现$\eta_j,\eta_{n-j+1}$. 从前面的关系式(4)可知：$\xi_j$(同样地，$\xi_{n-j+1}$)和$\xi_k(k\neq j,n-j+1)$正交. 而对于$1\leqslant j\leqslant m$，

$$(\xi_j,\xi_j)=\frac{1}{2}(\eta_j+\mathrm{i}\eta_{n-j+1},\eta_j+\mathrm{i}\eta_{n-j+1})$$

$$= \frac{1}{2}[(\eta_j,\eta_j) + 2\mathrm{i}(\eta_{n-j+1},\eta_j) - (\eta_{n-j+1},\eta_{n-j+1})]$$

$$= \frac{1}{2}[1 + 0 - 1] = 0,$$

$$(\xi_j,\xi_{n-j+1}) = \frac{1}{2}(\eta_j + \mathrm{i}\eta_{n-j+1},\eta_j - \mathrm{i}\eta_{n-j+1})$$

$$= \frac{1}{2}[(\eta_j,\eta_j) + \mathrm{i}(\eta_{n-j+1},\eta_j) - \mathrm{i}(\eta_j,\eta_{n-j+1})$$

$$+ (\eta_{n-j+1},\eta_{n-j+1})] = \frac{1}{2}[1 + 1] = 1,$$

$$(\xi_{n-j+1},\xi_{n-j+1}) = 0,$$

所以基 $\xi_1,\xi_2,\cdots,\xi_n$ 的度量矩阵为



2. $\dim V = n = 2m+1$.

此时令

$$\xi_j = \frac{1}{\sqrt{2}}(\eta_j + \mathrm{i}\eta_{n-j+1})$$

$$\xi_{n-j+1} = \frac{1}{\sqrt{2}}(\eta_j - \mathrm{i}\eta_{n-j+1}) \quad (j = 1,2,\cdots,m)$$

$$\xi_{m+1} = \eta_{m+1}.$$

不难验证，基 $\xi_1,\xi_2,\cdots,\xi_n$ 的度量矩阵也是

$$\begin{pmatrix} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \end{pmatrix}.$$

在复欧氏空间中，具有上述度量矩阵的基 $\xi_1,\xi_2,\cdots,\xi_n$ 称为**第二类规范基**.

复欧氏空间中的线性变换

我们现在讨论复欧氏空间中的几类重要线性变换.

定义 设 $\boldsymbol{A}$ 是 n 维复欧氏空间 V 的一个线性变换. 如果 V 的一个线性变换 $\boldsymbol{A}^*$ 满足如下条件：对一切 $\alpha,\beta\in V$，有

$$(\boldsymbol{A}\alpha,\beta)=(\alpha,\boldsymbol{A}^*\beta), \tag{5}$$

则称 $\boldsymbol{A}^*$ 为 $\boldsymbol{A}$ 的**共轭变换**.

现设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 为 V 的一组基，这组基的度量矩阵记为 $G=(g_{ij})$. 又设 $\boldsymbol{A}$ 与 $\boldsymbol{A}^*$ 在这组基下的矩阵分别为 $A=(a_{ij})$ 与 $B=(b_{ij})$. (5)式等价于

$$(\boldsymbol{A}\varepsilon_i,\varepsilon_j)=(\varepsilon_i,\boldsymbol{A}^*\varepsilon_j)\quad(i,j=1,2,\cdots,n).$$

此即

$$\sum_{k=1}^n g_{kj}a_{ki}=\sum_{k=1}^n g_{ik}b_{kj}.$$

写成矩阵形式，就是

$$A'G=GB. \tag{6}$$

G 是满秩复对称矩阵，故有

$$B=G^{-1}A'G.$$

由此即知：复欧氏空间内任一线性变换的共轭变换都是存在而且唯一的. 如果取 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 为 V 的第一类规范基，则 $G=E$，此时 $B=A'$.

定义 n 维复欧氏空间 V 内一个线性变换 $\boldsymbol{A}$，如果满足 $\boldsymbol{A}^*=\boldsymbol{A}$，则称为**对称变换**；如果满足 $\boldsymbol{A}^*=-\boldsymbol{A}$，则称为**反对称变换**.

在 V 内任取一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，设其度量矩阵为 G. 从(6)式即知：

1. $\boldsymbol{A}$ 是对称变换的充分必要条件是：$\boldsymbol{A}$ 在 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵 A 满足

$$A'G=GA.$$

当 $G=E$(第一类规范基)时，A 为对称矩阵；

2. $\boldsymbol{A}$ 是反对称变换的充分必要条件是：$\boldsymbol{A}$ 在 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵 A 满足

$$A'G+GA=0.$$

当 $G=E$ 时，A 为反对称矩阵.

定义 设 $\boldsymbol{R}$ 为 n 维复欧氏空间 V 内的一个线性变换. 如果 $\boldsymbol{R}$ 可逆，且 $\boldsymbol{R}^{-1}=\boldsymbol{R}^*$，则 $\boldsymbol{R}$ 称为 V 内的一个**正交变换**.

命题 3.3 n 维复欧氏空间 V 内的一个线性变换 $\boldsymbol{R}$ 是正交变换的充分必要条件是：对一切 $\alpha,\beta\in V$，有

$$(\boldsymbol{R}\alpha,\boldsymbol{R}\beta)=(\alpha,\beta).$$

证 （i）必要性. 若 $\boldsymbol{R}$ 是正交变换，则 $\boldsymbol{R}^{-1}=\boldsymbol{R}^*$，于是

$$(\boldsymbol{R}\alpha,\boldsymbol{R}\beta)=(\alpha,\boldsymbol{R}^*(\boldsymbol{R}\beta))=(\alpha,\boldsymbol{R}^{-1}(\boldsymbol{R}\beta))=(\alpha,\beta).$$

（ii）充分性. 在 V 内取一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$. 设其度量矩阵为 $G=(g_{ij})$. 又设 $\boldsymbol{R}$ 在此组基下的矩阵为 $R=(r_{ij})$. 因为

$$(\boldsymbol{R}\varepsilon_i,\boldsymbol{R}\varepsilon_j)=(\varepsilon_i,\varepsilon_j)=g_{ij},$$

故有

$$\sum_{k=1}^{n}\sum_{l=1}^{n}g_{kl}t_{ki}t_{lj}=g_{ij}.$$

写成矩阵形式，就是

$$R'GR=G.$$

G 是满秩的，由上式即知 R 可逆（否则 $R'GR$ 不满秩），于是 $\boldsymbol{R}$ 可逆. 又因为

$$R'G=GR^{-1},$$

与(6)式比较即知 $\boldsymbol{R}^*=\boldsymbol{R}^{-1}$. ▌

这个命题的证明过程同时指出，一个线性变换 $\boldsymbol{R}$ 是一个正交变换的充分必要条件是：它在某一组基下（等价地，在任一组基下）的矩阵 R 满足关系式

$$R'GR=G,$$

其中 G 是该基的度量矩阵.

下面我们来阐明复欧氏空间内正交变换与反对称变换之间的一个重要关系. 利用这个关系，可以在一定程度上把正交变换集合

的研究转化为反对称变换集合的研究.

定义 线性空间 V 内一个线性变换 $\boldsymbol{A}$ 如果不以 -1 作为特征值,则称为**非异线性变换**.

显然,一个线性变换 $\boldsymbol{A}$ 是非异的,其充分必要条件是:$\boldsymbol{E}+\boldsymbol{A}$ 没有零特征值,亦即 $\boldsymbol{E}+\boldsymbol{A}$ 可逆.

定理 4 设 V 是 n 维复欧氏空间.如果 $\boldsymbol{R}$ 是 V 内一个非异正交变换,则 $\boldsymbol{R}$ 可表成

$$\boldsymbol{R}=(\boldsymbol{E}-\boldsymbol{S})(\boldsymbol{E}+\boldsymbol{S})^{-1}=(\boldsymbol{E}+\boldsymbol{S})^{-1}(\boldsymbol{E}-\boldsymbol{S}), \tag{7}$$

其中 $\boldsymbol{S}$ 是一个非异的反对称变换.反之,如果 $\boldsymbol{S}$ 是 V 内一个非异的反对称变换,则由(7)式确定的 $\boldsymbol{R}$ 是一个非异正交变换.

证 在 V 内取一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,设其度量矩阵为 G.

(i) 如果 $\boldsymbol{R}$ 是一个非异正交变换,则 $\boldsymbol{E}+\boldsymbol{R}$ 可逆.令

$$\boldsymbol{S}=(\boldsymbol{E}-\boldsymbol{R})(\boldsymbol{E}+\boldsymbol{R})^{-1}=(\boldsymbol{E}+\boldsymbol{R})^{-1}(\boldsymbol{E}-\boldsymbol{R}). \tag{8}$$

因为

$$2\boldsymbol{E}-(\boldsymbol{E}+\boldsymbol{R})=\boldsymbol{E}-\boldsymbol{R},$$

两边右乘 $(\boldsymbol{E}+\boldsymbol{R})^{-1}$,得

$$2(\boldsymbol{E}+\boldsymbol{R})^{-1}-\boldsymbol{E}=(\boldsymbol{E}-\boldsymbol{R})(\boldsymbol{E}+\boldsymbol{R})^{-1}=\boldsymbol{S}.$$

于是 $\boldsymbol{E}+\boldsymbol{S}=2(\boldsymbol{E}+\boldsymbol{R})^{-1}$,即 $\boldsymbol{S}$ 是非异的,而且

$$\boldsymbol{E}+\boldsymbol{R}=2(\boldsymbol{E}+\boldsymbol{S})^{-1}.$$

移项,得

$$\begin{aligned}\boldsymbol{R}&=2(\boldsymbol{E}+\boldsymbol{S})^{-1}-\boldsymbol{E}=2\boldsymbol{E}(\boldsymbol{E}+\boldsymbol{S})^{-1}-(\boldsymbol{E}+\boldsymbol{S})(\boldsymbol{E}+\boldsymbol{S})^{-1}\\&=[2\boldsymbol{E}-(\boldsymbol{E}+\boldsymbol{S})](\boldsymbol{E}+\boldsymbol{S})^{-1}=(\boldsymbol{E}-\boldsymbol{S})(\boldsymbol{E}+\boldsymbol{S})^{-1}.\end{aligned}$$

再证 $\boldsymbol{S}$ 是一反对称变换.设 $\boldsymbol{R},\boldsymbol{S}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵分别为 R,S.因为 $\boldsymbol{R}$ 为正交变换,故有

$$R'GR=G. \tag{9}$$

从(8)式可知,有

$$(E+R)S=E-R.$$

两边取转置,得

$$S'(E+R')=E-R'.$$

以 GR 右乘上式两边，利用(9)式，得

$$S'G(R+E)=G(R-E).$$

两边再右乘$(R+E)^{-1}$，利用(8)式，即得

$$S'G=G(R-E)(R+E)^{-1}=-GS,$$

因而 $S'G+GS=0$，故 $\boldsymbol{S}$ 是一反对称变换.

（ii）反之，当 $\boldsymbol{S}$ 是非异反对称变换时，有

$$S'G+GS=0. \tag{10}$$

对由(7)式确定的 $\boldsymbol{R}$，有

$$(E+S)R=E-S.$$

两边转置，得

$$R'(E+S')=E-S'.$$

两边右乘 G，利用(10)式，有

$$R'G(E-S)=G(E+S).$$

两边再右乘$(E+S)^{-1}$，利用关系式

$$R=(E-S)(E+S)^{-1},$$

即得

$$R'GR=G.$$

这表明 $\boldsymbol{R}$ 是一正交变换，而且

$$\begin{aligned}\boldsymbol{E}+\boldsymbol{R}&=\boldsymbol{E}+(\boldsymbol{E}-\boldsymbol{S})(\boldsymbol{E}+\boldsymbol{S})^{-1}\\&=(\boldsymbol{E}+\boldsymbol{S})(\boldsymbol{E}+\boldsymbol{S})^{-1}+(\boldsymbol{E}-\boldsymbol{S})(\boldsymbol{E}+\boldsymbol{S})^{-1}\\&=2(\boldsymbol{E}+\boldsymbol{S})^{-1}.\end{aligned}$$

故 $\boldsymbol{R}$ 非异. ▌

习 题 三

1. 在 $\mathscr{R}^4$ 中定义内积如下：设

$$\alpha=(x_1,x_2,x_3,x_4),\ \beta=(y_1,y_2,y_3,y_4),$$

则令

$$(\alpha,\beta)=x_1y_1+x_2y_2+x_3y_3+x_4y_4.$$

证明：

（1）关于上述内积 $\mathscr{R}^4$ 成一准欧氏空间，并求其符号差；

（2）在 $\mathscr{R}^4$ 内取一组基

$$\varepsilon_1 = (1,0,0,0),\quad \varepsilon_2 = (0,1,0,0),$$

$$\varepsilon_3 = (0,0,1,0),\quad \varepsilon_4 = (0,0,0,1),$$

求内积在这组基下的度量矩阵.

2. 题 1 中所定义的准欧氏空间内的正交变换称为**广义洛仑兹**(Lorentz)**变换**. 设 $\boldsymbol{U}$ 是 $\mathscr{R}^4$ 内一个线性变换，在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$（参看题 1 的(2)）下的矩阵为 U，令

$$I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

证明：

（1）$\boldsymbol{U}$ 为广义洛仑兹变换的充分必要条件是：

$$U'IU = I;$$

（2）设 $\boldsymbol{U}$ 为广义洛仑兹变换，$U=(u_{ij})$. 证明：

$$|u_{44}| \geqslant 1.$$

3. 续上题. $\mathscr{R}^4$ 中一个向量 (x_1,x_2,x_3,x_4) 如果满足

$$x_1^2 + x_2^2 + x_3^2 - x_4^2 < 0,$$

则称为一个**类时向量**. 如果同时又有 $x_4>0$，则称为**正类时向量**. 又，$\mathscr{R}^4$ 内一个广义洛仑兹变换 $\boldsymbol{U}$ 如果满足：$u_{44}\geqslant 1$，则称为一个**洛仑兹变换**. 证明：一个广义洛仑兹变换 $\boldsymbol{U}$ 是洛仑兹变换的充分必要条件是：它把正类时向量变为正类时向量.

4. 续上题. 证明变换

$$\boldsymbol{S}(x_1,x_2,x_3,x_4) = (-x_1, -x_2, -x_3, x_4)$$

是一个洛仑兹变换.

5. 续上题. 行列式为 1 的洛仑兹变换为**正常洛仑兹变换**. 证明：任一非正常洛仑兹变换 $\boldsymbol{U}$ 都可以表作

$$U = SU_1,$$

其中U_1是正常洛仑兹变换.

6. 设 V 是 n 维复欧氏空间,$\alpha \in V$,如果对一切 $\beta \in V$ 都有 $(\alpha,\beta)=0$,证明 $\alpha=0$.

7. 设 V 是 n 维复欧氏空间,$\alpha \in V$,$\alpha \neq 0$. 令

$$M = \{\beta \in V \mid (\alpha,\beta) = 0\},$$

求 $\dim M$.

8. 设 V 是 n 维复欧氏空间,在 V 中找出一个非零子空间 M,使 M 关于 V 的内积不再是一个复欧氏空间.

9. 设 $\xi_1,\xi_2,\cdots,\xi_n$,是 n 维复欧氏空间 V 的一组第二类规范基,证明:如果线性变换 $\boldsymbol{A}$ 在这组基下的矩阵成若当块:

$$\begin{pmatrix} \lambda_0 & 1 & & \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0 \end{pmatrix},$$

则 $\boldsymbol{A}$ 是一个对称变换.

10. 证明复欧氏空间内共轭线性变换有如下关系式:

(1) $\boldsymbol{E}^* = \boldsymbol{E}$;　　(2) $(\boldsymbol{A}^*)^* = \boldsymbol{A}$;

(3) $(\lambda \boldsymbol{A})^* = \lambda \boldsymbol{A}^*$;　　(4) $(\boldsymbol{A}+\boldsymbol{B})^* = \boldsymbol{A}^* + \boldsymbol{B}^*$;

(5) $(\boldsymbol{AB})^* = \boldsymbol{B}^* \boldsymbol{A}^*$.

11. 设 $\boldsymbol{A}$ 是复欧氏空间 V 内的一个对称变换,证明:$\boldsymbol{A}$ 的属于不同特征值的特征向量互相正交.

12. 设 M 是 n 维复欧氏空间 V 的一个子空间,令

$$M^{\perp} = \{\alpha \in V \mid (\alpha,m) = 0,\ \text{对一切}\ m \in M\}.$$

问是否 $M \cap M^{\perp} = \{0\}$? 举例说明之.

13. 设 $\boldsymbol{A}$ 是 n 维复欧氏空间 V 内的一个线性变换,M 是 $\boldsymbol{A}$ 的不变子空间,证明:$M^{\perp}$ 是 $\boldsymbol{A}^*$ 的不变子空间.

14. 命 L 表示 n 维复欧氏空间 V 内全体反对称变换所成的集合,证明:

(1) L 关于线性变换的加法和数乘构成复数域上的线性空间；

(2) 求 $\dim L$；

(3) 设 $\boldsymbol{A},\boldsymbol{B}\in L$，证明$[\boldsymbol{A},\boldsymbol{B}]=\boldsymbol{AB}-\boldsymbol{BA}\in L$.

§4 满秩反对称双线性度量空间

定义 设 $f(\alpha,\beta)$是数域 K 上的线性空间 V 上的一个双线性函数．如果对一切 $\alpha,\beta\in V$ 都有

$$f(\alpha,\beta)=-f(\beta,\alpha),$$

则称 $f(\alpha,\beta)$是一个**反对称双线性函数**．

如果 V 是一个 n 维线性空间，那么，反对称双线性函数 $f(\alpha,\beta)$在任一组基下的矩阵 A 都是反对称矩阵：$A'=-A$. 于是有

$$|A|=|A'|=|-A|=(-1)^n|A|.$$

当 n 是奇数时，得到$|A|=-|A|$，故$|A|=0$. 这说明奇数维线性空间中的反对称双线性函数不可能是满秩的．

定义 设 V 是复数域 $\mathscr{C}$ 上 $n=2m$ 维线性空间，而 $f(\alpha,\beta)$是 V 上一个满秩反对称双线性函数．定义 V 内两个向量 α,β 的内积为

$$(\alpha,\beta)=f(\alpha,\beta),$$

称具有这种内积的线性空间为**辛空间**．

我们简单介绍一下辛空间的一些基本概念．

一、正交性

若$(\alpha,\beta)=0$，则称 α 与 β 正交．此时$(\beta,\alpha)=-(\alpha,\beta)=0$，故正交性具有对称的性质．显然，现在每个非零向量 α 都与自己正交：$(\alpha,\alpha)=0$.

二、基的度量矩阵

设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组基．令

$$(\varepsilon_i,\varepsilon_j) = g_{ij} \quad (i,\ j = 1,2,\cdots,n),$$

称 $G=(g_{ij})$ 为这组基的**度量矩阵**．若设

$$\alpha = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n,$$

$$\beta = y_1\varepsilon_1 + y_2\varepsilon_2 + \cdots + y_n\varepsilon_n,$$

则

$$(\alpha,\beta) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}x_i y_j = X'GY.$$

三、辛基

命题 4.1 设 V 是 $n=2m$ 维辛空间，则在 V 内存在一组基 $\eta_1,\eta_2,\cdots,\eta_n$，其度量矩阵为

$$G = \begin{pmatrix} A & & & \\ & A & & \\ & & \ddots & \\ & & & A \end{pmatrix}, \text{其中 } A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

这样的基称为**第一类辛基**．

证 对 m 作数学归纳法．

$m=1$ 时，$\dim V=2$. 在 V 内任取一组基 $\varepsilon_1,\varepsilon_2$. 由于内积满秩，反对称，故

$$(\varepsilon_1,\varepsilon_1) = 0, \quad (\varepsilon_1,\varepsilon_2) = k \neq 0, \quad (\varepsilon_2,\varepsilon_2) = 0.$$

令 $\eta_1=\dfrac{1}{\sqrt{k}}\varepsilon_1,\eta_2=\dfrac{1}{\sqrt{k}}\varepsilon_2$（$\sqrt{k}$ 为 k 的任一平方根）即可．

设命题对 $2(m-1)$ 维辛空间成立，证明它对 $2m$ 维辛空间也成立（此处设 $m\geqslant 2$）.

在 V 内任取一非零向量 ε_1. 因为内积满秩，必有 $\varepsilon_2\in V$，使 $(\varepsilon_1,\varepsilon_2)=k\neq 0$. 命

$$M = \{\alpha \in V \mid (\alpha,\varepsilon_i) = 0, i = 1,2\}.$$

M 显然为 V 的子空间．若 $\alpha\in M\cap L(\varepsilon_1,\varepsilon_2)$，则

$$\alpha = k_1\varepsilon_1 + k_2\varepsilon_2.$$

$$0 = (\alpha,\varepsilon_1) = k_2(\varepsilon_2,\varepsilon_1) \Rightarrow k_2 = 0,$$

$$0=(\alpha,\varepsilon_2)=k_1(\varepsilon_1,\varepsilon_2)\Rightarrow k_1=0.$$

故 $M\cap L(\varepsilon_1,\varepsilon_2)=\{0\}$. 另一方面 $\varepsilon_1,\varepsilon_2$ 线性无关(因若 $\varepsilon_2=l\varepsilon_1$,则 $(\varepsilon_2,\varepsilon_1)=l(\varepsilon_1,\varepsilon_1)=0$,与假设矛盾),把它们扩充成 V 的一组基

$$\varepsilon_1,\varepsilon_2,\varepsilon_3,\cdots,\varepsilon_n.$$

设

$$\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n.$$

由于齐次线性方程组

$$\begin{cases}(\varepsilon_1,\alpha)=x_1(\varepsilon_1,\varepsilon_1)+x_2(\varepsilon_1+\varepsilon_2)+\cdots+x_n(\varepsilon_1,\varepsilon_n)=0,\\(\varepsilon_2,\alpha)=x_1(\varepsilon_2,\varepsilon_1)+x_2(\varepsilon_2+\varepsilon_2)+\cdots+x_n(\varepsilon_2,\varepsilon_n)=0\end{cases}$$

的系数矩阵秩为 2,故 $\dim M=n-2$. 于是

$$V=L(\varepsilon_1,\varepsilon_2)\oplus M.$$

不难验证,M 关于 V 的内积是一个 $2(m-1)$ 维辛空间 . 按归纳假设,在 M 内存在一组基 $\eta_3,\eta_4,\cdots,\eta_n$,使内积在这组基下的度量矩阵为

$$\begin{pmatrix}0&1&&&\\-1&0&&&\\&&\ddots&&\\&&&0&1\\&&&-1&0\end{pmatrix}.$$

现在再令

$$\eta_1=\frac{1}{\sqrt{k}}\varepsilon_1,\ \eta_2=\frac{1}{\sqrt{k}}\varepsilon_2,$$

则 $\eta_1,\eta_2,\cdots,\eta_n$ 即为所求的基 . ▌

推论 设 V 是 $n=2m$ 维辛空间,则在 V 内存在一组基 $\xi_1,\xi_2,\cdots,\xi_n$,其度量矩阵为

$$\begin{pmatrix}0&E\\-E&0\end{pmatrix},$$

其中 E 为 m 阶单位矩阵 . 这种基称为**第二类辛基** .

证 设 $\eta_1,\eta_2,\cdots,\eta_n$ 为 V 的一组第一类辛基 . 令

$$\begin{aligned}\xi_i &= \eta_{2i-1} \\ \xi_{m+i} &= \eta_{2i}\end{aligned} \quad (i = 1,2,\cdots,m).$$

通过计算内积不难验证 $\xi_1,\xi_2,\cdots,\xi_n$ 即为所求的基. ▌

复欧氏空间一节中所论述的共轭变换、对称、反对称变换和正交变换的概念以及有关的命题,都可以平行地推移到辛空间中来.

定义 设 $\boldsymbol{A}$ 为 $n=2m$ 维辛空间 V 内的一个线性变换. 如果 V 内一个线性变换 $\boldsymbol{A}^*$ 满足如下条件:对一切 $\alpha,\beta\in V$,有

$$(\boldsymbol{A}\alpha,\beta) = (\alpha,\boldsymbol{A}^*\beta),$$

则称 $\boldsymbol{A}^*$ 为 $\boldsymbol{A}$ 的**共轭变换**.

在 V 内取一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,设其度量矩阵为 G. $\boldsymbol{A},\boldsymbol{A}^*$ 在这组基下的矩阵分别记为 A,B,则不难验证,有

$$A'G = GB.$$

G 可逆,故

$$B = G^{-1}A'G.$$

这说明,对 V 内任一线性变换 $\boldsymbol{A}$,其共轭变换 $\boldsymbol{A}^*$ 存在而且唯一.

如果 $\boldsymbol{A}^*=\boldsymbol{A}$,则 $\boldsymbol{A}$ 称为**对称变换**.

如果 $\boldsymbol{A}^*=-\boldsymbol{A}$,则 $\boldsymbol{A}$ 称为**反对称变换**.

如果 $\boldsymbol{R}$ 可逆,且 $\boldsymbol{R}^{-1}=\boldsymbol{R}^*$,则称 $\boldsymbol{R}$ 为**辛变换**. 一个线性变换 $\boldsymbol{R}$ 是**辛变换**的充分必要条件是

$$(\boldsymbol{R}\alpha,\boldsymbol{R}\beta) = (\alpha,\beta).$$

命题 4.2 在 V 内取一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,设其度量矩阵为 G. 又设线性变换 $\boldsymbol{A},\boldsymbol{R}$ 在这组基下的矩阵分别为 A,R,则有

(i) $\boldsymbol{A}$ 是对称变换的充分必要条件是

$$A'G = GA;$$

(ii) $\boldsymbol{A}$ 是反对称变换的充分必要条件是

$$A'G + GA = 0;$$

(iii) $\boldsymbol{R}$ 是辛变换的充分必要条件是

$$R'GR = G.$$

证明留给读者作为练习.

V 内一个线性变换 $\boldsymbol{A}$ 如不以 -1 为特征值，则称为**非异线性变换**．此时 $\boldsymbol{E}+\boldsymbol{A}$ 可逆．

定理 5 设 V 是 $n=2m$ 维辛空间．如果 $\boldsymbol{R}$ 是 V 内一个非异辛变换，则 $\boldsymbol{R}$ 可表成

$$\boldsymbol{R}=(\boldsymbol{E}-\boldsymbol{S})(\boldsymbol{E}+\boldsymbol{S})^{-1}=(\boldsymbol{E}+\boldsymbol{S})^{-1}(\boldsymbol{E}-\boldsymbol{S}), \tag{1}$$

其中 $\boldsymbol{S}$ 是一个非异的反对称变换．反之，如果 $\boldsymbol{S}$ 是 V 内一个非异的反对称变换，则由(1)式确定的 $\boldsymbol{R}$ 是一个非异辛变换．

证明留给读者作为练习．

设 $\boldsymbol{R}$ 为辛空间 V 内一个辛变换，又设 $\varepsilon_1,\eta_1,\varepsilon_2,\eta_2,\cdots,\varepsilon_m,\eta_m$ 为 V 内一组第一类辛基．此时其度量矩阵为

$$G=\begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & & & \\ & & \ddots & & \\ & & & 0 & 1 \\ & & & -1 & 0 \end{pmatrix}.$$

$\boldsymbol{R}$ 在此组基下的矩阵设为 R，则有

$$R'GR=G.$$

一个满足上述条件的 n 阶复方阵 R 称为一个 $2m$ 阶**辛矩阵**.

下面我们来介绍一类重要的辛变换．取定复数 c，又设 ε 为辛空间 V 内一非零向量，定义 V 内线性变换：

$$\boldsymbol{T}\alpha=\alpha+c(\alpha,\varepsilon)\varepsilon.$$

对任意 $\alpha,\beta\in V$，我们有

$$\begin{aligned}(\boldsymbol{T}\alpha,\boldsymbol{T}\beta)&=(\alpha+c(\alpha,\varepsilon)\varepsilon,\beta+c(\beta,\varepsilon)\varepsilon)\\&=(\alpha,\beta)+c(\alpha,\varepsilon)(\varepsilon,\beta)+c(\beta,\varepsilon)(\alpha,\varepsilon)\\&\quad+c^2(\alpha,\varepsilon)(\beta,\varepsilon)(\varepsilon,\varepsilon)\\&=(\alpha,\beta)-c(\alpha,\varepsilon)(\beta,\varepsilon)+c(\beta,\varepsilon)(\alpha,\varepsilon)=(\alpha,\beta).\end{aligned}$$

故 $\boldsymbol{T}$ 为 V 内一个辛变换，这种辛变换称为辛空间 V 的**辛平移**．

为了确切地描述辛平移，我们使用记号

$$\boldsymbol{T}(c,\varepsilon)\alpha=\alpha+c(\alpha,\varepsilon)\varepsilon$$

来表示辛平移.

命题 4.3 设 V 为 $2m$ 维辛空间, ε 为 V 内一非零向量, 则有

(i) 对任意复数 c_1, c_2 有

$$\boldsymbol{T}(c_1, \varepsilon)\boldsymbol{T}(c_2, \varepsilon) = \boldsymbol{T}(c_1 + c_2, \varepsilon).$$

(ii) 设 $\boldsymbol{R}$ 为 V 内任一辛变换, 那么

$$\boldsymbol{R}\boldsymbol{T}(c, \varepsilon)\boldsymbol{R}^{-1} = \boldsymbol{T}(c, \boldsymbol{R}(\varepsilon)).$$

(iii) 设 a 为非零复数, 我们有

$$\boldsymbol{T}(c, a\varepsilon) = \boldsymbol{T}(a^2 c, \varepsilon).$$

证 (i) 按辛平移的定义, 我们有

$$\begin{aligned}\boldsymbol{T}(c_1, \varepsilon)\boldsymbol{T}(c_2, \varepsilon)\alpha &= \boldsymbol{T}(c_1, \varepsilon)(\alpha + c_2(\alpha, \varepsilon)\varepsilon) \\ &= \alpha + c_1(\alpha, \varepsilon)\varepsilon + c_2(\alpha, \varepsilon)(\varepsilon + c_1(\varepsilon, \varepsilon)\varepsilon) \\ &= \alpha + (c_1 + c_2)(\alpha, \varepsilon)\varepsilon = \boldsymbol{T}(c_1 + c_2, \varepsilon)\alpha.\end{aligned}$$

(ii) 同样地, 我们有

$$\begin{aligned}\boldsymbol{R}\boldsymbol{T}(c, \varepsilon)\boldsymbol{R}^{-1}(\alpha) &= \boldsymbol{R}(\boldsymbol{R}^{-1}(\alpha) + c(\boldsymbol{R}^{-1}(\alpha), \varepsilon)\varepsilon) \\ &= \alpha + c(\boldsymbol{R}^{-1}(\alpha), \varepsilon)\boldsymbol{R}(\varepsilon) = \alpha + c(\alpha, \boldsymbol{R}(\varepsilon))\boldsymbol{R}(\varepsilon) \\ &= \boldsymbol{T}(c, \boldsymbol{R}(\varepsilon))(\alpha).\end{aligned}$$

(iii) 我们有

$$\begin{aligned}\boldsymbol{T}(c, a\varepsilon)(\alpha) &= \alpha + c(\alpha, a\varepsilon)(a\varepsilon) \\ &= \alpha + a^2 c(\alpha, \varepsilon)\varepsilon = \boldsymbol{T}(a^2 c, \varepsilon)(\alpha). \quad \blacksquare\end{aligned}$$

如果 $(\alpha, \varepsilon) = 0$, 那么 $\boldsymbol{T}(c, \varepsilon)\alpha = \alpha$. 特别地, 在辛空间中有 $(\varepsilon, \varepsilon) = 0$, 故 $\boldsymbol{T}(c, \varepsilon)\varepsilon = \varepsilon$.

习 题 四

1. 证明: 在 $n = 2m$ 维辛空间中存在一组基

$$\eta_1, \eta_2, \cdots, \eta_n,$$

使其度量矩阵为

$$G = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix},$$

其中 I 为 m 阶方阵，其形式为

$$I=\begin{pmatrix} & & 1\\ & \iddots & \\ 1 & & \end{pmatrix}.$$

2. 设 $\boldsymbol{R}$ 为 $n=2m$ 维辛空间 V 内的一个辛变换，如果 $\boldsymbol{R}$ 有 n 个互不相同的特征值，证明 V 内存在一组第一类辛基，使 $\boldsymbol{R}$ 在此组基下的矩阵成对角形．

（提示：若 $\boldsymbol{R}\alpha=\lambda\alpha,\boldsymbol{R}\beta=\mu\beta$，且 $(\alpha,\beta)\neq 0$，则有 $\mu=1/\lambda$.）

3. 设 V 为 $2m$ 维辛空间，证明 V 内两组第一类辛基间的过渡矩阵为辛矩阵．

4. 设 $\boldsymbol{R}$ 为 $2m$ 维辛空间 V 内的一个线性变换，证明下列命题互相等价：

(i) $\boldsymbol{R}$ 为 V 内辛变换；

(ii) $\boldsymbol{R}$ 把 V 内的第一类辛基变为第一类辛基；

(iii) $\boldsymbol{R}$ 在 V 的第一类辛基下的矩阵为辛矩阵.

习 题 解 答

第 一 章

习 题 一

1.(1)　$x_1=0, x_2=2, x_3=\frac{5}{3}, x_4=-\frac{4}{3}$.

(2)　$x_1=-\frac{1}{2}x_5, x_2=-1-\frac{1}{2}x_5$,

$x_3=0, x_4=-1-\frac{1}{2}x_5$.

(3)　无解.

(4)　$x_1=-8, x_2=3, x_3=6, x_4=0$.

(5)　$x_1=\frac{3}{17}x_3-\frac{13}{17}x_4$,　$x_2=\frac{19}{17}x_3-\frac{20}{17}x_4$.

(6)　无解.

(7)　$x_1=\frac{1}{6}+\frac{5}{6}x_4, x_2=\frac{1}{6}-\frac{7}{6}x_4, x_3=\frac{1}{6}+\frac{5}{6}x_4$.

3.(1)　无；　(2)　有；　(3)　有；　(4)　有.

5.(1)　当 $\lambda\neq 1, -2$ 时有唯一解：

$$x_1=-\frac{1+\lambda}{2+\lambda},\ x_2=\frac{1}{2+\lambda},\ x_3=\frac{(1+\lambda)^2}{2+\lambda};$$

当 $\lambda=1$ 时有解：$x_1=1-x_2-x_3$，x_2, x_3 任取；

当 $\lambda=-2$ 时无解.

(2)　当 $a\neq 1, b\neq 0$ 时有唯一解：

$$x_1=\frac{1-2b}{b(1-a)}, x_2=\frac{1}{b}, x_3=\frac{4b-2ab-1}{b(1-a)};$$

当 $a=1, b=\frac{1}{2}$ 时有解：$x_1=2-x_3, x_2=2$，x_3 任取；

当 $a=1, b\neq\frac{1}{2}$ 时无解；

当 $b=0$ 时无解．

（3） 当 $a=0, b=2$ 时有解．一般解为：

$x_1=-2+x_3+x_4+5x_5$， $x_2=3-2x_3-2x_4-6x_5$.

习 题 二

3. $\beta=6\alpha_1-3\alpha_2+\alpha_3+3\alpha_4+3\alpha_5$（解不唯一）.

4.（1） $\beta=\frac{5}{4}\alpha_1+\frac{1}{4}\alpha_2-\frac{1}{4}\alpha_3-\frac{1}{4}\alpha_4$.

（2） $\beta=\alpha_1-\alpha_3$.

5.（1） 线性无关； （2） 线性无关； （3） 线性相关；

（4） 线性相关．

14.（1） α_1,α_2 为一个极大线性无关部分组，秩$=2$.

（2） $\alpha_1,\alpha_2,\alpha_3$ 为一个极大线性无关部分组，秩$=3$.

（3） α_1,α_3 为一个极大线性无关部分组，秩$=2$.

习 题 三

1.（1） 4； （2） 3； （3） 2； （4） 3； （5） 5.

2.（1） $\alpha_2,\alpha_3,\alpha_4$ 为一个极大线性无关部分组，秩$=3$.

（2） $\alpha_1,\alpha_2,\alpha_4$ 为一个极大线性无关部分组，秩$=3$.

5. 秩$=n$.

习 题 四

1.（1） $\eta_1=(1,-2,1,0,0)$， $\eta_2=(1,-2,0,1,0)$，

$\eta_3=(5,-6,0,0,1)$.

（2） $\eta_1=(18,36,24,39,9)$.

（3） $\eta_1=(-1,-1,1,2,0), \eta_2=(1,0,0,5,4)$.

（4） 仅有零解．

（5） $\eta_1=(0,1,2,1)$.

(6) $\eta_1=(0,1,1,0,0)$, $\eta_2=(0,1,0,1,0)$,

$\eta_3=(1,-5,0,0,3)$.

习 题 五

1. (1) $\gamma_0=\left(\frac{2}{3},\frac{1}{6},0,0,0\right)$;

$\eta_1=(0,1,2,0,0)$, $\eta_2=(0,-1,0,2,0)$,

$\eta_3=(2,5,0,0,6)$.

$\eta=\gamma_0+k_1\eta_1+k_2\eta_2+k_3\eta_3$.

(2) $\gamma_0=\left(\frac{2}{3},-\frac{1}{6},\frac{1}{6},0\right)$; $\eta_1=(-2,-7,-5,6)$.

$\eta=\gamma_0+k_1\eta_1$.

第 二 章

习 题 一

3. (2) $\begin{pmatrix}3 & -2\\ 4 & 8\end{pmatrix}$. (3) $\begin{pmatrix}1 & n\\ 0 & 1\end{pmatrix}$.

(4) $\begin{pmatrix}\cos n\varphi & -\sin n\varphi\\ \sin n\varphi & \cos n\varphi\end{pmatrix}$. (5) (0).

(6) $\begin{pmatrix}2 & 3 & -1\\ -2 & -3 & 1\\ -2 & -3 & 1\end{pmatrix}$.

(7) $a_{11}x^2+a_{22}y^2+2a_{12}xy+2b_1x+2b_2y+c$.

(8) 当 $n=2k$ 时为： 当 $n=2k+1$ 时为：

$4^k\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$; $4^k\begin{pmatrix}1 & -1 & -1 & -1\\ -1 & 1 & -1 & -1\\ -1 & -1 & 1 & -1\\ -1 & -1 & -1 & 1\end{pmatrix}$.

(9) $\begin{pmatrix} \lambda^n & n\lambda^{n-1} & \frac{n(n-1)}{2}\lambda^{n-2} \\ 0 & \lambda^n & n\lambda^{n-1} \\ 0 & 0 & \lambda^n \end{pmatrix}$.

5. $f(A)=\begin{pmatrix} 5 & 1 & 3 \\ 8 & 0 & 3 \\ -2 & 1 & -2 \end{pmatrix}$.

6. $\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ 0 & x_{11} & x_{12} \\ 0 & 0 & x_{11} \end{pmatrix}$，其中 x_{11}, x_{12}, x_{13} 为任意数．

习 题 二

2.（1） $P=\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$．（2） $Q=\begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

5. $\begin{pmatrix} a & 0 \\ 0 & \frac{1}{a} \end{pmatrix}$

$$=\begin{pmatrix} 1 & a-1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & \frac{1}{a} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -a & 1 \end{pmatrix}.$$

习 题 三

3.（1） $A^{-1}=\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

（2） $A^{-1}=\frac{1}{3}\begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & -2 \\ -3 & 2 & -1 \end{pmatrix}$.

（3） $A^{-1}=\begin{pmatrix} 1 & -4 & -3 \\ 1 & -5 & -3 \\ -1 & 6 & 4 \end{pmatrix}$.

(4) $A^{-1}=\begin{pmatrix} 22 & -6 & -26 & 17 \\ -17 & 5 & 20 & -13 \\ -1 & 0 & 2 & -1 \\ 4 & -1 & -5 & 3 \end{pmatrix}$.

(5) $A^{-1}=\frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$.

(6) $A^{-1}=\begin{pmatrix} -7 & 5 & 12 & -19 \\ 3 & -2 & -5 & 8 \\ 41 & -30 & -69 & 111 \\ -59 & 43 & 99 & -159 \end{pmatrix}$.

(7) $A^{-1}=\begin{pmatrix} 1 & -3 & 11 & -38 \\ 0 & 1 & -2 & 7 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

(8) $A^{-1}=\begin{pmatrix} 2 & -1 & 0 & 0 \\ -3 & 2 & 0 & 0 \\ -5 & 7 & -3 & -4 \\ 2 & -2 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.

(9) $A^{-1}=\frac{1}{6}\begin{pmatrix} -1 & 3 & -7 & 20 \\ -7 & -3 & 5 & -10 \\ 9 & 3 & -3 & 6 \\ 3 & 3 & -3 & 6 \end{pmatrix}$.

(10) $A^{-1}=\frac{1}{2^5}\begin{pmatrix} 16 & -8 & 4 & -2 & 1 \\ 0 & 16 & -8 & 4 & -2 \\ 0 & 0 & 16 & -8 & 4 \\ 0 & 0 & 0 & 16 & -8 \\ 0 & 0 & 0 & 0 & 16 \end{pmatrix}$.

4. (1) $X=\begin{pmatrix}2 & -23\\ 0 & 8\end{pmatrix}$.　　(2) $X=\frac{1}{6}\begin{pmatrix}11 & 3 & 6\\ -1 & -3 & 0\\ 4 & 6 & 0\end{pmatrix}$.

(3) $X=\frac{1}{6}\begin{pmatrix}-2 & 2 & 8\\ 4 & 2 & 2\\ 4 & 5 & 8\end{pmatrix}$.

(4) $X=\begin{pmatrix}1 & -1 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0\\ 1 & 1 & -1 & -1 & 0 & \cdots & 0 & 0 & 0 & 0\\ 0 & 1 & 1 & -1 & -1 & \cdots & 0 & 0 & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 1 & -1\\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 2\end{pmatrix}_{n\times n}$.

5. $A^{-1}=\begin{pmatrix}0 & 0 & \cdots & 0 & a_n^{-1}\\ a_1^{-1} & 0 & \cdots & 0 & 0\\ 0 & a_2^{-1} & \cdots & 0 & 0\\ \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & \cdots & a_{n-1}^{-1} & 0\end{pmatrix}$.

习　题　四

1. (2) 写成分块形成：

$$\begin{pmatrix}A_1 & & \\ & A_2 & \\ & & A_3\end{pmatrix}^3=\begin{pmatrix}A_1^3 & & \\ & A_2^3 & \\ & & A_3^3\end{pmatrix}.$$

其中

$$A_1=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}, A_2=(2), A_3=\begin{pmatrix}-3 & 1 & 0\\ 0 & -3 & 1\\ 0 & 0 & -3\end{pmatrix}.$$

(3) 写成分块形式：

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}^5 = \begin{pmatrix} A_1^5 & 0 \\ 0 & A_2^5 \end{pmatrix} = 0,$$

其中

$$A_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

2. 写成分块形式：

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}^{-1} = \begin{pmatrix} A_1^{-1} & 0 \\ 0 & A_2^{-1} \end{pmatrix},$$

其中

$$A_1 = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{pmatrix}.$$

3. $X^{-1} = \begin{pmatrix} 0 & C^{-1} \\ A^{-1} & 0 \end{pmatrix}$.

4. $D^{-1} = \begin{pmatrix} A^{-1} & 0 \\ -B^{-1}CA^{-1} & B^{-1} \end{pmatrix}$.

第 三 章

习 题 一

7. (1) -6； (2) 25； (3) -483； (4) 0；

(5) 24； (6) $\frac{3}{8}$.

8. (1) $-2(x^3+y^3)$； (2) x^2y^2； (3) 0； (4) 48；

(5) 160.

习 题 二

8. $$A^*=\begin{pmatrix}-2&0&2\\0&0&1\\-6&2&5\end{pmatrix}.$$

9. (1) $A^*=\begin{pmatrix}0&1&1\\0&1&-2\\-3&2&-1\end{pmatrix}$, $A^{-1}=\frac{1}{3}\begin{pmatrix}0&1&1\\0&1&-2\\-3&2&-1\end{pmatrix}$.

(2) $A^*=\begin{pmatrix}-1&4&3\\-1&5&3\\1&-6&-4\end{pmatrix}$, $A^{-1}=\begin{pmatrix}1&-4&-3\\1&-5&-3\\-1&6&4\end{pmatrix}$.

(3) $A^*=\begin{pmatrix}1&0&-1&0\\0&1&0&0\\0&0&-1&1\\0&0&0&-1\end{pmatrix}$, $A^{-1}=A^*$.

16. (1) $x_1=1,x_2=1,x_3=1,x_4=1$.

(2) $x_1=1,x_2=2,x_3=-1,x_4=-2$.

习 题 三

1. $N(23145)=2$,偶排列;

$N(985467321)=31$,奇排列;

$N(375149)=6$,偶排列;

$N(n(n-1)(n-2)\cdots 321)=\dfrac{n(n-1)}{2}$,当 $n=4k,4k+1$ 时为偶排列,当 $n=4k+2,4k+3$ 时为奇排列;

$N((2n+1)(2n-1)\cdots 531)=\dfrac{n(n+1)}{2}$,当 $n=4k,4k+3$ 时为偶排列,当 $n=4k+1,4k+2$ 时为奇排列.

6. x^3 系数为 -1,x^4 系数为 2.

第 四 章

习 题 一

1. $M\cup N=\{-1,3,2,1,0,7,9,-3\}$.

   $M\cap N=\{-1,0,2\}$.

3. $M\cap N=\varnothing$.

4. $M\cap N=M$.

8. (1) 是； (2) 是； (3) 是； (4) 否； (5) 是；
   (6) 否； (7) 是.

习 题 二

1. (1) 线性无关，秩$=2$.
   (2) 线性相关，秩$=2$.
   (3) 线性无关，秩$=2$.
   (4) 线性无关，秩$=2$.
   (5) 线性无关，秩$=n$.

3. (1) 线性相关，秩$=1$.
   (2) 线性相关，秩$=2$.
   (3) 线性相关，秩$=2$.

4. $\dim\mathscr{Q}(\omega)=2$；$1,\omega$ 为一组基.

6. (1) 数域 K 上全体对称矩阵组成$\frac{n(n+1)}{2}$维线性空间. E_{ii} $(i=1,2,\cdots,n)$，$E_{ij}+E_{ji}$ $(i<j)$为其一组基（其中 E_{ij} 为 i 行 j 列处为 1，其余元素为零的 n 阶方阵）.
   (2) 数域 K 上全体主对角线元素之和为零的 n 阶方阵组成 n^2-1 维线性空间，$E_{11}-E_{ii}$ $(i=2,3,\cdots,n)$，E_{ij} $(i\neq j)$为其一组基.
   (3) 该线性空间维数为 1，任一不等于 1 的正实数都可作

为它的一组基.

(4) 该线性空间维数为 3,E,A,A^2 为一组基.

7. (1) $\beta=\frac{5}{4}\varepsilon_1+\frac{1}{4}\varepsilon_2-\frac{1}{4}\varepsilon_3-\frac{1}{4}\varepsilon_4$.

(2) $\beta=2\varepsilon_1+\varepsilon_2-3\varepsilon_3+2\varepsilon_4$.

9. 过渡矩阵 T 的第 $k+1$ 列($k=0,1,2,\cdots,m-1$)是(自上而下):

$(-1)^k a^k, (-1)^{k-1}C_k^{k-1}a^{k-1}, (-1)^{k-2}C_k^{k-2}, a^{k-2}, \cdots,$

$-C_k^1 a, 1, 0, \cdots, 0.$

10. (1) 过渡矩阵

$$T=\begin{pmatrix} 2 & 0 & 5 & 6 \\ 1 & 3 & 3 & 6 \\ -1 & 1 & 2 & 1 \\ 1 & 0 & 1 & 3 \end{pmatrix}.$$

$\beta=x_1\eta_1+x_2\eta_2+x_3\eta_3+x_4\eta_4$,其中

$$\begin{cases} x_1=\frac{4}{9}b_1+\frac{1}{3}b_2-b_3-\frac{11}{9}b_4, \\ x_2=\frac{1}{27}b_1+\frac{4}{9}b_2-\frac{1}{3}b_3-\frac{23}{27}b_4, \\ x_3=\frac{1}{3}b_1-\frac{2}{3}b_4, \\ x_4=-\frac{7}{27}b_1-\frac{1}{9}b_2+\frac{1}{3}b_3+\frac{26}{27}b_4. \end{cases}$$

(2) 过渡矩阵

$$T=\begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

$$\beta=\frac{3}{13}\varepsilon_1+\frac{5}{13}\varepsilon_2-\frac{2}{13}\varepsilon_3-\frac{3}{13}\varepsilon_4.$$

(3) 过渡矩阵

$$T=\frac{1}{4}\begin{pmatrix}3&7&2&-1\\1&-1&2&3\\-1&3&0&-1\\1&-1&0&-1\end{pmatrix}.$$

$$\beta=-2\eta_1-\frac{1}{2}\eta_2+4\eta_3-\frac{2}{3}\eta_4.$$

11. $\xi=(-a,-a,-a,a)\ (a\neq 0)$.

习 题 三

4.(3) $\dim C(A)=n$，$E_{ii}(i=1,2,\cdots,n)$为一组基.

5. $\dim C(A)=5$. 它的一组基为

$$\begin{pmatrix}1&0&0\\0&0&0\\-3&0&0\end{pmatrix},\begin{pmatrix}0&1&0\\0&0&0\\9&0&3\end{pmatrix},\begin{pmatrix}0&0&0\\1&0&0\\-1&0&0\end{pmatrix},\begin{pmatrix}0&0&0\\0&1&0\\3&0&1\end{pmatrix},$$

$$\begin{pmatrix}0&0&0\\0&0&0\\3&1&1\end{pmatrix}.$$

6. 该子空间维数为 4.

10.(1) 3 维，$\alpha_2,\alpha_3,\alpha_4$ 为一组基；

(2) 2 维，α_1,α_2 为一组基.

11. 2 维，它的一组基为：

$\eta_1=(-1,24,9,0)$，$\eta_2=(2,-21,0,9)$.

12.(1) 和的维数=3；$\alpha_1,\alpha_2,\beta_1$ 为一组基；

交的维数=1；$4\alpha_2-\alpha_1$ 为一组基.

(2) 和的维数=4；$\alpha_1,\alpha_2,\beta_1,\beta_2$ 为一组基；

交的维数=0.

(3) 和的维数=4；$\alpha_1,\alpha_2,\alpha_3,\beta_2$ 为一组基；

交的维数=1；β_1 为一组基.

第　五　章

习　题　一

1. (1)　当 $\alpha\neq0$ 时不是；　(2)　当 $a\neq0$ 时不是；

(3)　不是；(4)　是；(5)　是；(6)　是；

(7)　不是；(8)　是.

习　题　二

1. $\boldsymbol{A}\varepsilon_1=(1,-2,2,0)$，$\boldsymbol{A}\varepsilon_2=(0,-1,0,3)$，

$\boldsymbol{A}\varepsilon_3=(-1,0,2,3)$，$\boldsymbol{A}\varepsilon_4=(-2,-1,-1,3)$.

2. (1) $\begin{pmatrix}2&-1&0\\0&1&1\\1&0&0\end{pmatrix}$；

(2) $A=\begin{pmatrix}\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}\end{pmatrix}$，$B=\begin{pmatrix}0&0\\0&1\end{pmatrix}$，$AB=\begin{pmatrix}0&\frac{1}{2}\\0&\frac{1}{2}\end{pmatrix}$；

(3) $\begin{pmatrix}0&1&&\\&0&\ddots&\\&&\ddots&1\\&&&0\end{pmatrix}$；(4) $\begin{pmatrix}a&b&1&0&0&0\\-b&a&0&1&0&0\\0&0&a&b&1&0\\0&0&-b&a&0&1\\0&0&0&0&a&b\\0&0&0&0&-b&a\end{pmatrix}$；

(5) $\begin{pmatrix}-1&1&-2\\2&2&0\\3&0&2\end{pmatrix}$；　(6) $\frac{1}{7}\begin{pmatrix}-5&20&-20\\-4&-5&-2\\27&18&24\end{pmatrix}$；

(7) $\begin{pmatrix}2&3&-5\\-1&0&-1\\-1&1&0\end{pmatrix}$.

3. (2) $\begin{pmatrix} 0 & -c & b & 0 \\ -b & a-d & 0 & b \\ c & 0 & d-a & -c \\ 0 & c & -b & 0 \end{pmatrix}$.

4. A.

5. (1) $\begin{pmatrix} a_{33} & a_{32} & a_{31} \\ a_{23} & a_{22} & a_{21} \\ a_{13} & a_{12} & a_{11} \end{pmatrix}$; (2) $\begin{pmatrix} a_{11} & ka_{12} & a_{13} \\ \dfrac{a_{21}}{k} & a_{22} & \dfrac{a_{23}}{k} \\ a_{31} & ka_{32} & a_{33} \end{pmatrix}$;

(3) $\begin{pmatrix} a_{11}+a_{12} & a_{12} & a_{13} \\ a_{21}+a_{22}-a_{11}-a_{12} & a_{22}-a_{12} & a_{23}-a_{13} \\ a_{31}+a_{32} & a_{32} & a_{33} \end{pmatrix}$.

8. $\begin{pmatrix} 2 & -3 & 3 & 2 \\ \dfrac{2}{3} & -\dfrac{4}{3} & \dfrac{10}{3} & \dfrac{10}{3} \\ \dfrac{8}{3} & -\dfrac{16}{3} & \dfrac{40}{3} & \dfrac{40}{3} \\ 0 & 1 & -7 & -8 \end{pmatrix}$.

9. (1) $\begin{pmatrix} -1 & 2 & 4 & 0 \\ 0 & 2 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & -3 & -3 & 0 \end{pmatrix}$;

(2) $\dfrac{1}{4}\begin{pmatrix} 4 & 4 & -2 & -8 \\ 2 & 12 & -1 & -6 \\ 2 & -6 & 10 & 14 \\ 0 & 14 & -5 & -12 \end{pmatrix}$.

10. (1) $\begin{pmatrix} -2 & -\frac{3}{2} & \frac{3}{2} \\ 1 & \frac{3}{2} & \frac{3}{2} \\ 1 & \frac{1}{2} & -\frac{5}{2} \end{pmatrix}$.　(2) $\begin{pmatrix} -2 & -\frac{3}{2} & \frac{3}{2} \\ 1 & \frac{3}{2} & \frac{3}{2} \\ 1 & \frac{1}{2} & -\frac{5}{2} \end{pmatrix}$.

习　题　三

5. (1)　$\lambda_1=7$，　$V_{\lambda_1}=L(\varepsilon_1+\varepsilon_2)$；

　　$\lambda_2=-2$，　$V_{\lambda_2}=L(-4\varepsilon_1+5\varepsilon_2)$.

(2)　当 $a\neq0$ 时：$\lambda_1=a\mathrm{i}$，$V_{\lambda_1}=L(\varepsilon_1+\mathrm{i}\varepsilon_2)$；

　　$\lambda_2=-a\mathrm{i}$，$V_{\lambda_2}=L(\varepsilon_1+\mathrm{i}\varepsilon_2)$.

(3)　$\lambda_1=2$，　$V_{\lambda_1}=L(-2\varepsilon_1+\varepsilon_2)$；

　　$\lambda_2=1+\sqrt{3}$，$V_{\lambda_2}=(-3\varepsilon_1+\varepsilon_2+(\sqrt{3}-2)\varepsilon_3)$；

　　$\lambda_3=1-\sqrt{3}$，$V_{\lambda_2}=L(-3\varepsilon_1+\varepsilon_2-(\sqrt{3}+2)\varepsilon_3)$.

(4)　$\lambda_1=1$，$V_{\lambda_1}=L(\varepsilon_2,\varepsilon_1+\varepsilon_3)$；

　　$\lambda_2=-1$，$V_{\lambda_2}=L(\varepsilon_1-\varepsilon_3)$.

(5)　$\lambda_1=0$，　$V_{\lambda_1}=L(-3\varepsilon_1+\varepsilon_2-2\varepsilon_3)$；

　　$\lambda_2=\sqrt{14}\mathrm{i}$，

　　$V_{\lambda_2}=L((3+2\sqrt{14}\mathrm{i})\varepsilon_1+13\varepsilon_2+(2-3\sqrt{14}\mathrm{i})\varepsilon_3)$；

　　$\lambda_3=\sqrt{14}\mathrm{i}$，

　　$V_{\lambda_2}=L((3-2\sqrt{14}\mathrm{i})\varepsilon_1+13\varepsilon_2+(2+3\sqrt{14}\mathrm{i})\varepsilon_3)$.

(6)　$\lambda_1=1$，　$V_{\lambda_1}=L(3\varepsilon_1-6\varepsilon_2+20\varepsilon_3)$；

　　$\lambda_2=-2$，$V_{\lambda_2}=L(\varepsilon_3)$.

(7)　$\lambda_1=2$，$V_{\lambda_1}=L(\varepsilon_1+\varepsilon_2,\varepsilon_1+\varepsilon_3,\varepsilon_1+\varepsilon_4)$；

　　$\lambda_2=-2$，$V_{\lambda_2}=L(-\varepsilon_1+\varepsilon_2+\varepsilon_3+\varepsilon_4)$.

习 题 四

8. (1) $$\begin{pmatrix} -1 & -2 & 0 \\ 0 & -1 & 0 \\ 1 & 2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 0 \\ -2 & -2 & -1 \end{pmatrix} \begin{pmatrix} -1 & -2 & 0 \\ 0 & -1 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

(2) $$\begin{pmatrix} 0 & 8 & 0 \\ 1 & 6 & 0 \\ 0 & -4 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 3 & 0 & 8 \\ 3 & -1 & 6 \\ -2 & 0 & -5 \end{pmatrix} \begin{pmatrix} 0 & 8 & 0 \\ 1 & 6 & 0 \\ 0 & -4 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

(3) $$\begin{pmatrix} 2 & 5 & 0 \\ -2 & -3 & 1 \\ -2 & -1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 4 & 5 & -2 \\ -2 & -2 & 1 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 5 & 0 \\ -2 & -3 & 1 \\ -2 & -1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

9. (1) 做变数替换

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2-\sqrt{2} & 2+\sqrt{2} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix},$$

方程组化作：

$$\begin{cases} \dfrac{dz_1}{dx} = (3+\sqrt{2})z_1, \\ \dfrac{dz_2}{dx} = (3-\sqrt{2})z_2. \end{cases}$$

（2） 作变数替换

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 4 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix},$$

方程组化作：

$$\begin{cases} \dfrac{\mathrm{d}z_1}{\mathrm{d}x} = z_1, \\ \dfrac{\mathrm{d}z_2}{\mathrm{d}x} = -z_2, \\ \dfrac{\mathrm{d}z_3}{\mathrm{d}x} = 4z_3. \end{cases}$$

第 六 章

习 题 一

8.（2） $$A = \begin{pmatrix} -4 & -14 & 15 & 6 \\ 15 & -2 & -2 & -7 \\ -12 & -10 & 1 & 14 \\ 3 & 4 & -15 & -2 \end{pmatrix}.$$

（3） $$A = \begin{pmatrix} -9 & -29 & 11 & 35 \\ 25 & -3 & 69 & -11 \\ 1 & -123 & -3 & -11 \\ -5 & -1 & -1 & -9 \end{pmatrix}.$$

9.（2） $$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

（3） 过渡矩阵

$$T=\begin{pmatrix}1 & 1 & 0 & 0\\0 & 0 & 1 & 1\\0 & 0 & 1 & -1\\1 & -1 & 0 & 0\end{pmatrix}.$$

$f(A,B)$的矩阵为：

$$A=\begin{pmatrix}2 & 0 & 0 & 0\\0 & 2 & 0 & 0\\0 & 0 & 2 & 0\\0 & 0 & 0 & -2\end{pmatrix}.$$

（4） 秩＝4.

12.（2） $A=\begin{pmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & 1 & 0\\0 & 0 & 0 & -1\end{pmatrix}$. （3） $=\begin{pmatrix}5 & 2 & 10 & 14\\2 & 10 & 11 & 19\\10 & 11 & 37 & 47\\14 & 19 & 47 & 64\end{pmatrix}$.

（5） $\alpha=(0,0,1,1)$.

习 题 二

2.（1） $\begin{pmatrix}-1 & 2 & 0 & -2\\2 & -3 & 0 & 1\\0 & 0 & 0 & 0\\-2 & 1 & 0 & 2\end{pmatrix}$.

$$f(\alpha,\alpha)=-x_1^2+4x_1x_2-4x_1x_4-3x_2^2+2x_2x_4+2x_4^2.$$

（2） $\begin{pmatrix}1 & 0 & 0 & 0\\0 & -1 & 0 & 0\\0 & 0 & -1 & 0\\0 & 0 & 0 & 0\end{pmatrix}$. $f(\alpha,\alpha)=x_1^2-x_2^2-x_3^2$.

3.（1） $f(\alpha,\beta)=x_1y_1+x_1y_2+x_2y_1+x_2y_2+x_2y_3+x_2y_2$
$+x_3y_3+x_3y_4+x_4y_3+x_4y_4$.

（2） $f(\alpha,\beta)=\frac{1}{2}[x_1y_2+x_1y_3+x_1y_4+x_2y_1+x_2y_3+x_2y_4$

$+x_3y_1+x_3y_2+x_3y_4+x_4y_1+x_4y_2+x_4y_3]$.

4.(1) $-4y_1^2+4y_2^2+y_3^2$. (2) $y_1^2+y_2^2$;

(3) $y_1^2-y_2^2$; (4) $8y_1^2+2y_2^2-2y_3^2-8y_4^2$;

(5) $6y_1^2-y_2^2-2y_3^2-y_4^2$; (6) $y_1^2+2y_2^2-2y_3^2$;

(7) $y_1^2+y_2^2+2y_3^2-2y_4^2$;

(8) $y_1^2-y_2^2+y_3^2-y_4^2+\cdots+y_{2n-1}^2-y_{2n}^2$;

(9) 当 $n=2k+1$ 时：$y_1^2-y_2^2+y_3^2-y_4^2+\cdots+y_{2k-1}^2-y_{2k}^2$;

当 $n=2k$ 时： $y_1^2-y_2^2+y_3^2-y_4^2+\cdots+y_{2k-1}^2-y_{2k}^2$.

习 题 三

1.(1) $z_1^2+z_2^2+z_3^2$; (2) $z_1^2+z_2^2+z_3^2$;

(3) $z_1^2+z_2^2+z_3^2$; (4) $z_1^2+z_2^2$;

2.(1) $z_1^2+z_2^2-z_3^2$; (2) $z_1^2-z_2^2-z_3^2$;

(3) $z_1^2-z_2^2$; (4) $z_1^2+z_2^2$.

习 题 四

1.(1) 是； (2) 不是； (3) 是； (4) 是.

2.(1) $-\frac{4}{5}<t<0$; (2) t 取任何值二次型均非正定.

第 七 章

习 题 一

1.(1) $\frac{\pi}{2}$. (2) $\frac{\pi}{4}$. (3) $\arccos\frac{3}{\sqrt{77}}$.

6. $\pm\frac{1}{\sqrt{26}}(-4,0,-1,3)$.

10. $\eta_1=\frac{1}{\sqrt{2}}(\varepsilon_1+\varepsilon_5)$,

$$\eta_2=\frac{1}{\sqrt{10}}(\varepsilon_1-2\varepsilon_2+2\varepsilon_4-\varepsilon_5),$$

$$\eta_3=\frac{1}{2}(\varepsilon_1+\varepsilon_2+\varepsilon_3-\varepsilon_5).$$

11. $$\eta_1=\frac{1}{\sqrt{2}}(0,1,1,0,0),$$

$$\eta_2=\frac{1}{\sqrt{10}}(-2,1,-1,2,0),$$

$$\eta_3=\frac{1}{\sqrt{315}}(7,-6,6,13,5).$$

14. (1) $$\eta_1=\left(\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}},\frac{-1}{\sqrt{6}},0\right),$$

$$\eta_2=\left(\frac{4}{\sqrt{30}},\frac{-1}{\sqrt{30}},\frac{2}{\sqrt{30}},\frac{3}{\sqrt{30}}\right),$$

$$\eta_3=\left(\frac{-36}{\sqrt{5670}},\frac{39}{\sqrt{5670}},\frac{42}{\sqrt{5670}},\frac{33}{\sqrt{5670}}\right).$$

(2) $$\eta_1=\left(\frac{2}{\sqrt{6}},\frac{1}{\sqrt{6}},0,\frac{1}{\sqrt{6}}\right),$$

$$\eta_2=\left(\frac{-2}{\sqrt{30}},\frac{1}{\sqrt{30}},\frac{4}{\sqrt{30}},\frac{3}{\sqrt{30}}\right),$$

$$\eta_3=\left(\frac{-2}{\sqrt{21}},\frac{2}{\sqrt{21}},\frac{-3}{\sqrt{21}},\frac{2}{\sqrt{21}}\right).$$

17. $$T=\begin{pmatrix}\frac{1}{3} & -\frac{2}{\sqrt{6}} & 0 & \frac{\sqrt{2}}{3}\\ -\frac{2}{3} & 0 & \frac{1}{\sqrt{3}} & \frac{\sqrt{2}}{3}\\ 0 & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} & \frac{\sqrt{2}}{2}\\ \frac{2}{3} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{\sqrt{2}}{6}\end{pmatrix}.$$

习 题 三

5. (1) $T=\begin{pmatrix} \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{pmatrix}$.

(2) $T=\begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{2}{3\sqrt{5}} & -\frac{1}{3} \\ \frac{1}{\sqrt{5}} & \frac{4}{3\sqrt{5}} & -\frac{2}{3} \\ 0 & \frac{5}{3\sqrt{5}} & \frac{2}{3} \end{pmatrix}$.

(3) $T=\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.

(4) $T=\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{\sqrt{3}}{6} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{\sqrt{3}}{6} & -\frac{1}{2} \\ 0 & \frac{2}{\sqrt{6}} & \frac{\sqrt{3}}{6} & \frac{1}{2} \\ 0 & 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$.

(5) $$T=\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{\sqrt{3}}{6} & \frac{1}{2} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{\sqrt{3}}{6} & \frac{1}{2} \\ 0 & -\frac{2}{\sqrt{6}} & \frac{\sqrt{3}}{6} & \frac{1}{2} \\ 0 & 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}.$$

6. (1) $$\begin{cases} x_1=\frac{2}{3}y_1-\frac{2}{3}y_2-\frac{1}{3}y_3, \\ x_2=\frac{2}{3}y_1+\frac{1}{3}y_2+\frac{2}{3}y_3, \\ x_3=\frac{1}{3}y_1+\frac{2}{3}y_2-\frac{2}{3}y_3. \end{cases}$$

标准形：$-y_1^2+2y_2^2+5y_3^2$.

(2) $$\begin{cases} x_1=-\frac{2}{5}\sqrt{5}\,y_1+\frac{2\sqrt{5}}{15}y_2-\frac{1}{3}y_3, \\ x_2=\frac{1}{5}\sqrt{5}\,y_1+\frac{4\sqrt{5}}{15}y_2-\frac{2}{3}y_3, \\ x_3=\frac{\sqrt{5}}{3}y_2+\frac{2}{3}y_3. \end{cases}$$

标准形：$2y_1^2+2y_2^2-7y_3^2$.

(3) $$\begin{cases} x_1=\frac{1}{\sqrt{2}}y_2+\frac{1}{\sqrt{2}}y_4, \\ x_2=\frac{1}{\sqrt{2}}y_2-\frac{1}{\sqrt{2}}y_4, \\ x_3=\frac{1}{\sqrt{2}}y_1+\frac{1}{\sqrt{2}}y_3, \\ x_4=\frac{1}{\sqrt{2}}y_1-\frac{1}{\sqrt{2}}y_3. \end{cases}$$

标准形：$y_1^2+y_2^2-y_3^2-y_4^2$.

(4) $$\begin{cases} x_1=\frac{1}{2}y_1+\frac{1}{2}y_2-\frac{1}{2}y_3-\frac{1}{2}y_4, \\ x_2=\frac{1}{2}y_1-\frac{1}{2}y_2-\frac{1}{2}y_3+\frac{1}{2}y_4, \\ x_3=\frac{1}{2}y_1-\frac{1}{2}y_2+\frac{1}{2}y_3-\frac{1}{2}y_4, \\ x_4=\frac{1}{2}y_1+\frac{1}{2}y_2+\frac{1}{2}y_3+\frac{1}{2}y_4. \end{cases}$$

标准形：$y_1^2-3y_2^2-y_3^2+7y_4^2$.